OM nucleic - nucleic search, using sw model

GenCore version 5.1.3

Run on: February 1, 2003, 22:42:09 ; Search time 79 Seconds
                              (without alignments)
                              8200.505 Million cell updates/sec

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 ttttttttttttgagatggag..............ttaaacaaagcttagaagca 1442

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published_Applications_NA:*

```
 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1442	100.0	1442	9	US-10-146-130-1	Sequence 1, Appli
2	1442	100.0	1442	10	US-09-964-666-1	Sequence 1, Appli
3	1442	100.0	1442	10	US-09-964-412-1	Sequence 1, Appli
4	1442	100.0	1442	10	US-09-964-666-4	Sequence 4, Appli
5	1223.4	84.8	1418	10	US-09-964-412-4	Sequence 4, Appli
6	1223.4	84.8	1418	10	US-09-964-666-3	Sequence 3, Appli
7	1080.2	74.9	1381	10	US-09-964-412-3	Sequence 3, Appli
8	1080.2	74.9	1381	9	US-09-964-531-180	Sequence 180, Appl
9	587.8	40.8	65608	10	US-09-962-436-292	Sequence 292, Appl	c
10	587.8	40.8	65608	10	US-09-962-832-119	Sequence 119, App	c
11	524.4	36.4	62944	10	US-09-954-456-2257	Sequence 2257, Ap	c
12	513	35.6	99014	10	US-09-764-847-1157	Sequence 1157, Ap
13	490.4	34.0	21470	10	US-09-880-107-3428	Sequence 3428, Ap
14	479.2	33.2	92139	10	US-09-918-686-2	Sequence 2, Appli
15	479.2	33.2	51719	10	US-09-918-686-1	Sequence 1, Appl
16	475.4	33.0	14796	10	US-09-954-456-973	Sequence 973, Ap
17	475.4	33.0	14796	10	US-09-954-456-1636	Sequence 1636, Ap
18	475.4	33.0	14796	10	US-09-918-186A-3	Sequence 3, Appli
19	475.4	33.0	14796	10	US-09-880-107-3421	Sequence 3421, Ap
20	475.4	33.0	26657	10	US-09-810-673A-3	Sequence 3, Appli
21	471.2	32.7	12542	10	US-09-764-864-1774	Sequence 1774, Ap
22	467	32.4	15041	10	US-09-764-869-1421	Sequence 1421, Ap
23	464.8	32.2	57130	10	US-09-835-081-3	Sequence 3, Appli
24	447.8	31.1	14176	10	US-09-764-864-1644	Sequence 1644, Ap
25	440.8	30.6	13224	10	US-09-764-853-897	Sequence 897, App
26	440.2	30.5	110096	10	US-09-880-107-1542	Sequence 1542, Ap
27	439.2	30.4	110096	10	US-09-880-107-1542	Sequence 1542, Ap	c
28	433.2	30.0	20247	10	US-09-764-877-2680	Sequence 2680, Ap	c
29	429.8	29.8	27148	10	US-09-764-860-1046	Sequence 1046, Ap	c
30	426	29.5	32152	10	US-09-764-855-328	Sequence 328, App	c
31	426	29.5	32204	10	US-09-764-855-327	Sequence 327, App	c
32	425.6	29.5	24533	9	US-09-764-868-1349	Sequence 1349, Ap
33	425.6	29.5	34641	10	US-09-954-456-1110	Sequence 1110, Ap
34	425.6	29.5	34641	10	US-09-954-456-1787	Sequence 1787, Ap
35	421	29.2	65464	10	US-09-859-888-3	Sequence 3, Appli
36	420.6	29.2	2109	10	US-09-764-855-325	Sequence 325, App
37	420.6	29.2	17200	10	US-09-764-877-3390	Sequence 3390, Ap
38	419.2	29.1	48436	10	US-09-927-602-38	Sequence 38, Appli
39	419.2	29.1	6439	9	US-09-790-852-8	Sequence 8, Appli
40	419.2	29.1	75270	9	US-09-790-852-8	Sequence 8, Appli	c
41	418.4	29.0	21606	9	US-09-790-852-8	Sequence 1733, Ap	c
42	418.4	29.0	133893	9	US-10-161-510-1	Sequence 1733, Ap	c
43	417.2	28.9	38374	10	US-09-880-107-3463	Sequence 3463, Ap	c
44	414.4	28.7	7233	10	US-09-764-869-2123	Sequence 2123, Ap	c
45	412.8	28.6	17792	10	US-09-764-869-1599	Sequence 1599, Ap	c

ALIGNMENTS

RESULT 1
US-10-146-130-1

; Sequence 1, Application US/10146130
; Publication No. US20030004107A1
; GENERAL INFORMATION:
; APPLICANT: AVERBACK, PAUL
; TITLE OF INVENTION: METHOD OF PREVENTING CELL DEATH USING SEGMENTS OF
; TITLE OF INVENTION: NEURAL THREAD PROTEINS
; FILE REFERENCE: 59003.000007
; CURRENT APPLICATION NUMBER: US/10/146,130
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 1442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(1139)

US-10-146-130-1

Query Match 100.0%; Score 1442; DB 9; Length 1442;
Best Local Similarity 100.0%;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     1 TTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTGTGCCAGGCTGCAATGGCGCAA 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTGTGCCAGGCTGCAATGGCGCAA 60

Qy    61 TCTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCGCTCCC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCGCTCCC 120

Qy   121 CAGTAGCTGGGATTACAGGCACGCCACCACGCTCGGCTAATTTTGTATTTTTTAG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTAGCTGGGATTACAGGCACGCCACCACGCTCGGCTAATTTTGTATTTTTTAG 180

Qy   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
```

```
Qy    241 CCTCCGTCTCTCGGCCTCCCAAAGTGCTGTAGATACAGAGACTGGCCACCATGCCGGGCTGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCTCTCGGCCTCCCAAAGTGCTGTAGATACAGAGACTGGCCACCATGCCGGGCTGCC 300

Qy    301 TGGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGAGATACAGGACTGGCCAAGCTGGTCTCCTGAGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGAGATACAGGACTGGCCAAGCTGGTCTCCTGAGC 360

Qy    361 TCAAGCAGTCCACTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAGTCCACTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC 420

Qy    421 CTGGCCTTTTTATTTTTTATTTTTTTAAGACAGGGTGTCCCACTGTTACCCAGGATGAAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGCCTTTTTATTTTTTATTTTTTTAAGACAGGGTGTCCCACTGTTACCCAGGATGAAG 480

Qy    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTCGAACTCCTGAGATCAAGCATCCTCCCA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTCGAACTCCTGAGATCAAGCATCCTCCCA 540

Qy    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600

Qy    601 TTTTTATTTTTTATTTTTGAGACAGGAGTCTCACCCAGGCTGGAGTGCAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTATTTTTTATTTTTGAGACAGGAGTCTCACCCAGGCTGGAGTGCAGT 660

Qy    661 GGGCGAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGGCGAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCCA 720

Qy    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCGCCACCACGCCCAGCTAATTTTTGTATTTTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCGCCACCACGCCCAGCTAATTTTTGTATTTTT 780

Qy    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGGTCTCGATCTCTGGACCTTGTGATC 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGGTCTCGATCTCTGGACCTTGTGATC 840

Qy    841 TGCCTGCCTCCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCTTA 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCTCCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCTTA 900

Qy    901 TTTTTAATTTTTTGTTTGTTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTTAATTTTTTGTTTGTTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960

Qy    961 GGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGTCTCA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGTCTCA 1020

Qy   1021 GCCTCCCAAGCAGCTGGGAAACTGCACCACCACACCCCGCCTAATTTTTGTATTTT 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCAGCTGGGAAACTGCACCACCACACCCCGCCTAATTTTTGTATTTT 1080

Qy   1081 TCATTAGAGGCGGGGTTTCACCATATTGTCTCAAAACTCCTGACCTCAGGT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTGTCTCAAAACTCCTGACCTCAGGT 1140

Qy   1141 GACCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200

Qy   1201 CCGGCTAATTTAGATAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCCCAGGCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGCTAATTTAGATAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCCCAGGCT 1260

Qy   1261 GGTCTCAAACTTCTGGGCTTCAAGCGAATCCTTCCAAATGAGCCACAACACACCCAGCTCA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCAAACTTCTGGGCTTCAAGCGAATCCTTCCAAATGAGCCACAACACACCCAGCTCA 1320

Qy   1321 CATTTTTTAAACAGTTACATCTTTATTTAGTATACTAGAAAGTAATACAATAAAACATGT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTTAAACAGTTACATCTTTATTTAGTATACTAGAAAGTAATACAATAAAACATGT 1380

Qy   1381 CAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTAATCTTTTAAACTTTTAAACAAAGCTTTAGAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTAATCTTTTAAACTTTTAAACAAAGCTTTAGAG 1440

Qy   1441 CA 1442
          ||
Db   1441 CA 1442
```

RESULT 2
US-09-964-666-1
; Sequence 1, Application US/09964666
; Patent No. US20020104108A1
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
            Wands, Jack R.
TITLE OF INVENTION: Transgenic Animals and Cell Lines for
     Screening Drugs Effective for the Treatment or Preventio
     of Alzheimer's Disease
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
     ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
     STREET: 1100 New York Ave., Suite 600
     CITY: Washington
     STATE: DC
     COUNTRY: USA
     ZIP: 20005-3934
COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/09,964,666
     FILING DATE: 28-Sep-2001
     CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
     NAME: Esmond, Robert W.
     REGISTRATION NUMBER: 32,893
     REFERENCE/DOCKET NUMBER: 0609.4370000
TELECOMMUNICATION INFORMATION:
     TELEPHONE: 202-371-2600
     TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
     LENGTH: 1442 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: double
     TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
     NAME/KEY: CDS
     LOCATION: 15..1139
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-964-666-1

Query Match 100.0%; Score 1442; DB 10; Length 1442;
Best Local Similarity 100.0%; Pred. No. 7.2e-293;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TTTTTTTTTTTTTTGAGATGGAGTTTTCGGCTCTTGTTGCCCAGGCTGGAGTGCAATGCGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTTTGAGATGGAGTTTTCGGCTCTTGTTGCCCAGGCTGGAGTGCAATGCGCAA 60

Qy     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy    121 CAGTAGCTGGGATTACAGGCATGTGCACCACCACGCCTCGGCTAATTTTTGTATTTTTTTAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   121 CAGTAGCTGGGATTACAGGCGTGCACCCACGCTCGGCGTAATTTTTGTATTTTTTTTAG 180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   181 TAGAGATGGGAGTTTCTCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
Db   181 TAGAGATGGGAGTTTCTCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
Qy   241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCCTCTGCC 300
Db   241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCCTCTGCC 300
Qy   301 TGGCTAATTTTTGTGTAGAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTCCTGAGC 360
Db   301 TGGCTAATTTTTGTGTAGAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTCCTGAGC 360
Qy   361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCACCCGTGC 420
Db   361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCACCCGTGC 420
Qy   421 CTGGCCTTTTATTTTATTTTTTTTAAGACAGGGTGTCCCACTCTTACCCAGGATGAAG 480
Db   421 CTGGCCTTTTATTTTATTTTTTTTAAGACAGGGTGTCCCACTCTTACCCAGGATGAAG 480
Qy   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTGAACTCCTGAGATCAAGCATCCTCCTG 540
Db   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTGAACTCCTGAGATCAAGCATCCTCCTG 540
Qy   541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600
Db   541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600
Qy   601 TTTTTTATTTTTATTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 660
Db   601 TTTTTTATTTTTATTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 660
Qy   661 GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCCCA 720
Db   661 GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCCCA 720
Qy   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCCTCCCGGGTTAATTTTTT 780
Db   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCCTCCCGGGTTAATTTTTT 780
Qy   781 TAGTAGAGAATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCTGGTGTGATC 840
Db   781 TAGTAGAGAATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCTGGTGTGATC 840
Qy   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCGCCGGCTTA 900
Db   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCGCCGGCTTA 900
Qy   901 TTTTTAATTTTTGTTTGTTTGAAATGGGAAGTCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
Db   901 TTTTTAATTTTTGTTTGTTTGAAATGGGAAGTCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
Qy   961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020
Db   961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020
Qy  1021 GCCTCCCAAGCAGCTGGGATTACGGGCGCACCTGCCACCACACCCGCTAATTTTTGTATTT 1080
Db  1021 GCCTCCCAAGCAGCTGGGATTACGGGCGCACCTGCCACCACACCCGCTAATTTTTGTATTT 1080
Qy  1081 TCATTAGAGGGGGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCTCAGT 1140
Db  1081 TCATTAGAGGGGGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCTCAGT 1140
Qy  1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200
Db  1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAG 1200
Qy  1201 CCGGCTCAATTTAGATAAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCCCAGGCT 1260
Db  1201 CCGGCTCAATTTAGATAAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCCCAGGCT 1260
```

```
Qy  1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACACCCAGCCAGTCA 1320
Db  1261 GGTCTCAAACTTCTGGCTTCATGCAATCCTTCCAAATGAGCCACCAACACCCAGCCAGTCA 1320
Qy  1321 CATTTTTTTAAACAGTTACACATCTTTTAATTTTTAGTATACTACTAGAAAGTAATACAATAAACATGT 1380
Db  1321 CATTTTTTTAAACAGTTACACATCTTTTAATTTTTAGTATACTACTAGAAAGTAATACAATAAACATGT 1380
Qy  1381 CAAACCTGCAAAATTCAGTAGTACAGAGAGTTTCTTTTTATAAACTTTTAAACAAAAGCTTTAGAG 1440
Db  1381 CAAACCTGCAAAATTCAGTAGTACAGAGAGTTTCTTTTTATAAACTTTTAAACAAAAGCTTTAGAG 1440
Qy  1441 CA 1442
         ||
Db  1441 CA 1442
```

```
Query Match             100.0%;  Score 1442;  DB 10;  Length 1442;
Best Local Similarity   100.0%;  Pred. No. 7.2e-293;
Matches 1442;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     1 TTTTTTTTTTTTTGAGATGGAGTTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTTTTTTTTTTGAGATGGAGTTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
```

```
Qy    61 TCTCAGCTCACCGCAACCTCCGGCTCCCGGGTTCAAGCGATTCTCCTGCCTC         120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTCAGCTCACCGCAACCTCCGGCTCCCGGGTTCAAGCGATTCTCTGCCTC          120

Qy   121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTGTATTTTTTTAG    180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTGTATTTTTTTAG    180

Qy   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATC    240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC    240

Qy   241 CCTCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCAGATGTC    300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTCGAACCTCCAGATGATC     300

Qy   301 TGGCTAATTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC     360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGCTAATTTTGTGGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGC    360

Qy   361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC    420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC    420

Qy   421 CTGGCCTTTTTATTTATTTTATTTTTTTTAAGACACAGGGTCCCACTCTTACCCAGGATGAAG    480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGCCTTTTTATTTATTTTATTTTTTTTAAGACACAGGGTCCCACTCTTACCAGGATGAAG    480

Qy   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCCTCCTG    540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCCTCCTG    540

Qy   541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACCACTGAATTTTTTA     600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACCACTGAATTTTTTA     600

Qy   601 TTTTTATTTTATTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT      660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTTTATTTTATTTAATTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT     660

Qy   661 GGCGCAAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTATTATTCTCCTGCCCCA     720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GGCGCAAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTATTATTCTCCTGCCCCA    720

Qy   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTGTATTTT     780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTGTATTTT    780

Qy   781 TAGTAGAGATGGGGTTTCACCAGTTGATCTTGATCTTGATCTCGGACCTTGTGATC      840
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TAGTAGAGATGGGGTTCACCATGTTGGTCAGGCTGGTCTCGGACCTTGTGATC       840

Qy   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCGCCTTA    900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCGCCTTA    900

Qy   901 TTTTTTAATTTTTTGTTTTTGTTTTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT    960
         |||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTTTTAATTTTTTGTTTTTGTTTTTTGAAATGGAATCTCACTCTGTTACCAGGCTGGAGTGCAAT    960

Qy   961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA    1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA    1020

Qy  1021 GCCTCCCAAGCAGCTGGGATTACAGGCACACACCACCACCCCTTAATTTTTGTATTT     1080
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GCCTCCCAAGCAGCTGGGATTACAGGCGCACACCACCACCCCGGCTAATTTTTGTATTT     1080

Qy  1081 TCATTTAGAGACGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT    1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TCATTTAGAGACGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT    1140

Qy  1141 GACCCACCTGCCTCGAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCACCCAG    1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCCACCTGCCTCGAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCACCCAG    1200

Qy  1201 CCGGCTAATTTAGTAGATAAAAAAATATGTAGCAATGTGGGGGGGGTCTTTGCTATGTTGCCCAGGCT    1260
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CCGGCTAATTTAGTAGATAAAAAAATATGTAGCAATGTGGGGGGGGTCTTTGCTATGTTGCCCAGGCT    1260

Qy  1261 GGTCTCAAACTTCTGGCTTCATGCGAATCCTTCCAAAATGAGCCACAACACCAGCCAGTCA    1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GGTCTCAAACTTCTGGCTTCATGCGAATCCTTCCAAAATGAGCCACAACACCAGCCAGTCA    1320

Qy  1321 CATTTTTTTAAACAGTTACATCTTTATTTTAGTTATTACTGAAAAGTAAATACAATAAACATGT    1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CATTTTTTTAAACAGTTACATCTTTATTTTAGTTATTACTAGAAAAGTAATACAATAAACATGT    1380

Qy  1381 CAAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTATAAACTTTTATAATAAACTTTTAGAG    1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTATAACTTTTATAATAAACTTTTAGAG    1440

Qy  1441 CA 1442
         ||
Db  1441 CA 1442
```

```
Query Match          84.8%;  Score 1223.4;  DB 10;  Length 1418;
Best Local Similarity  96.9%;  Pred. No. 3.5e-247;
Matches 1375;  Conservative  0;  Mismatches 31;  Indels 13;  Gaps 12;

Qy    2 TTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT 61
```

RESULT 5
US-09-964-412-4
; Sequence 4, Application US/09964412
; Patent No. US20020129391A1
; GENERAL INFORMATION:
;   APPLICANT: de la Monte, Suzanne
;   Wands, Jack R.
;   TITLE OF INVENTION: Transgenic Animals and Cell Lines for
;   Screening Drugs Effective for the Treatment or Preventio
;   of Alzheimer's Disease
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
;   STREET: 1100 New York Ave., Suite 600
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20005-3934
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/964,412
;   FILING DATE: 28-Sep-2001
;   CLASSIFICATION: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Esmond, Robert W.
;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 0609.4370000
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-371-2600
;   TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1418 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: both
;   MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-964-412-4

Query Match   84.8%; Score 1223.4; DB 10; Length 1418;
Best Local Similarity 96.9%; Pred. No. 3.5e-247;
Matches 1375; Conservative 0; Mismatches 31; Indels 13; Gaps 12;

```
Qy      1 TTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 CTCAGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC  121

Qy     61 CTCAGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTA-GCTGGGATTACAGGCGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTTAGTAG  180

Qy    121 AGTAGCTGGGATTACAGGCGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTTAGTAG  179
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATC  240

Qy    181 TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 CCTCCGCTCGGCCTCCCAAAGTGCTAGGATTACAGGACTGAGCCACCATGCCCGGCCTGC  298

Qy    241 CCTCCGCTCGGCCTCCCAAAGTGCTAGGATTACAGGACTGAGCCACCATGCCCGGCCTGC  299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 CTGGGCTAATTTTGTGTGTAGAAACAGGGTTTCACTGAGTGTCCCAAGCTGGTCTCCTGAG  358

Qy    300 CTGGGCTAATTTTGTGTGTAGAAACAGGGTTTCACTGAGTGTCCCAAGCTGGTCTCCTGAG  359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 CTCAAGCAGTCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTG  418

Qy    360 CTCAAGCAGTCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTG  419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCACTCTTTACCCAGGATGAA  478

Qy    420 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCACTCTTTACCCAGGATGAA  479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCATCCTTCCT  537

Qy    480 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCTGAGATCAAGCATCCTTCCT  539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 GCCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACATTACCACCACTACACCTGGCTAATTT  597

Qy    540 GCCTCAGCCTCCC-AAGTAGCTGGGACCAAAGACATTACCACCACTACACCTGGCTAATTT  598
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TATTTTTATTTTTTAATTTTTTTGAGACAGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCA  657

Qy    599 TATTTTTATTTTTTAATTTTTTTGAGACAGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCA  658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 GTGGCGCAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCC  717

Qy    659 GTGGCGCAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCC  718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCAGGCCCACCTAATTTTTTGTATT  777

Qy    719 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCAGGCCCACCTAATTTTTTGTATT  778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    778 TTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGGTCTCGAATCTCTTGACCTTGT  837

Qy    779 TTTAGTAGAGATGGGG-TTCACCATGTTGGCCAGGTTGGTCTCGAATCTCTTGACCTTGT  836
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    838 GATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG-CGTGAGCCACCACGCCCG  897

Qy    837 GATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG-CGTGAGCCACCACGCCCG  895
          ||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db    898 GCCTATTTTTAATTTTTTGTTTGTTTGTTTTGAAATGGGAATCTCACTCTGTTACCCAGGCTGGAGT  957

Qy    896 GCCTATTTTTAATTTTTTGTTTGTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGT  955
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    958 GCAATGGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTG  1017

Qy    956 GCAATGGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTG  1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 TCTCAGCCTCCCAAGAGCTGGGATTACGGGCACCACCCGGCCTG-CACCACACCCGCTAATTTTTG  1076

Qy   1016 TCTCAGCCTCCCAAGAGCTGGGATTACGGGCACCACCCGGCCTG-CACCACACCCGCTAATTTTTG  1075
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1076 TATTTTCATTAGAGAGCGGGGGGTTTCACCAATATTTGTCAGGCTGGTCTCAAACTCCTGACCT  1135

Qy   1076 TATTTTCATTAGAGAGCGGGGGGTTTCACCAATATTTGTCAGGCTGGTCTCAAACTCCTGACCT  1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 TATTTTCATTAGAGAGCGGGGGGTTTCACCAATATTTGTCAGGCTGGTCTCAAACTCCTGACCT  1136

Qy   1136 CAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCA  1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1137 CAGGTGACCCACCTGCCTCAGCCTTCCAAAAGTGCTGGGATTACAGGCGTGA--CGCCTCA  1194

Qy   1196 CCCAGCCGCGGCTAATTTAGTAAAAAAAATGTAGCAAGTGGGGGGGTCTTGCTATGTTGCCC  1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1196 CCCAGCCGCGGCTAATTTAGTAAAAAAAAATGTAGCAAGTGGGGGGGTCTTGCTATGTTGCCC  1255

Qy   1256 AGGCTGGTCTCAAACTCCTGGGCTTCATGCGAAATCCTTCCCAAAATAGCTGGGACCACCAACAAACCAGCC  1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 AGGCTGGTCTCAAACTTCGGGCTTCATGCGATCCTTCATTAATACTAGAAAGTGAACAAACACCCAAGC  1314

Qy   1316 AGTCACATTTTTTAAAACAGTTACATCTTTATTTATTTTAGTATACTAGGAAGTAATACAATAA  1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 AGTCACA-TTTTAAACAGTTACAACTTTTATTTATTTTAGTATACTAGAAAAGTGAATAACAGATAAC  1373

Qy   1376 CATGTCAAACCTGCAAATTCAGTAGTAGTAACACAGAGTTCTTT  1414
          |||||||||||||||||||||||||||||||||||||||||||
Db   1374 ATGGCGGAACCTGCAAATTCGAGTAGTAGTACAGAGTCTTTT  1412
```

```
RESULT 6
US-09-964-666-3
; Sequence 3, Application US/09964666
; Patent No. US2002010410Sa1
; GENERAL INFORMATION:
;  APPLICANT: de la Monte, Suzanne
;       Wands, Jack R.
;  TITLE OF INVENTION: Transgenic Animals and Cell Lines for
;       Screening Drugs Effective for the Treatment or Preventio
;       of Alzheimer's Disease
;  NUMBER OF SEQUENCES: 14
;  CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.C.
;   STREET: 1100 New York Ave., Suite 600
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20005-3934
;  COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;  CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/964,666
;   FILING DATE: 28-Sep-2001
;   CLASSIFICATION: <Unknown>
;  ATTORNEY/AGENT INFORMATION:
;   NAME: Esmond, Robert W.
;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 0609.4370000
;  TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-371-2600
;   TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 3:
;  SEQUENCE CHARACTERISTICS:
;   LENGTH: 1381 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: both
;  MOLECULE TYPE: cDNA
;  SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-964-666-3

Query Match             74.9%;  Score 1080.2;  DB 10;  Length 1381;
Best Local Similarity   94.4%;                 Pred. No. 3e-217;
Matches 1317;  Conservative    0;  Mismatches   53;  Indels   25;  Gaps   18;

Qy      2 TTTTTTTTTTTGAGATGGAGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGGCGCAAT 61
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db      1 TTTTTTTTTTTTGAGATGGAGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGGGGCAAT 60

Qy     62 CTCAGCTCACCGCAACCTCCGGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 121
          ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGCAACCTCCGCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy    122 AGTA-GCTGGGATTACAGGGCATGTGCACCCA-CCACGCTCGGCTAATTTTTGTATTTTTTTAG 180
          |||| ||||||||||||||| || || || || || |||||||||||||||||||||||||||||
Db    121 AGTAGGCTGGGATTACAGGCCATTGCCA--CCACCACGCTCGGCTAATTTTTGTATTTTTTTAG 179

Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGATGATC 240
          |||||||||| |||||||| |||||||||||||||||||||||||||||||||||||| |||
Db    180 TAGAGGATGGGAGTTTCTCCAATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATC 238

Qy    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATTACAGGCGTGTCAGCCACCATGCCCGG-CTCTGC 299
          |||| || ||||||||||||||||||||||||| |||||| ||| |||||||||||| ||||||
Db    239 CTCCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCATGCCCGGCCTCTGC 298

Qy    300 CTGGCTAATTTTTGTGTTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAG 359
          ||||||||||| |||||| ||||||||||||| ||||| || ||||||||||||| |||||||
Db    299 CTGGCTAATTTTTGTGTTAGTAGAAACAGGGTTTCACTGATGTGCCAAAGCTGGTCTCCTGAG 358

Qy    360 CTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTG 419
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||||| |||
Db    359 CTCAAGCAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCCGTG 418

Qy    420 CCTGGCCTTTTTTATTTTTATTTTTTTTTAAGACAACAGGGTCTCCACTCTTACCCAGGATGAA 479
          |||||||||||| ||||||| |||||||| ||||||||||||||||||||| ||||||||| ||
Db    419 CCTGGCCTTTTTTATTTTTATTTTTTTTTAAGACAGGGTCTCCACTCTGTACCCAGGATGAA 478

Qy    480 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAATCCTCCT 539
          |||||||||| ||||||||||| |||||||||||||||||||||||||||| ||||||||||
Db    479 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAATCCTCCT 537

Qy    540 GCCTCAGCCTCCC-AAGTAGCTGGGACTACAGGCAAGACATGCACCACTACACGTGGCTAATTT 598
          |||||||||||| |||||||||||||||||| |||| |||||||||||||| |||||||||||
Db    538 GCCTCAGCCTCCCAAAGTAGCTGGGACTACAGGCAAGACATGCCACCACTACACCTGGCTAATTT 597

Qy    599 TATTTTTTATTTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCA 658
          ||||||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db    598 TATTTTTTTATTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCA 657

Qy    659 GTGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCC 718
          |||||| |||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    658 GTGGCGCAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCC 717

Qy    719 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTGTATT 778
          |||||||||||||||||||||||||||| |||||||||||||||| |||||||||| ||||
Db    718 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCAGCTAATTTTTGTATT 777

Qy    779 TTTAGTAGAGATGGGG-TTCACCATGTTCGGCCAGGTTGAT-CTTGATCTCCTGGACCTTGT 836
          ||||||||||||||||| || |||||||| ||||||| |||| |||||||||||||||||||||
Db    778 TTTAGTAGAGATGGGGGTTTCACCATGTTGGTCCAGGATGATCTCGATCTCCTGACCTTGT 836

Qy    837 GATCTGCCTGCCTCGGCCTCCCAAACTGCTGGGATTACAGG-CGTGAGCCACCACGCCCG 895
          ||||||||||||||| ||||||||||||||||||||||||| |||||||||||||| ||||
Db    837 GATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCG 895

Qy    896 GCTTATTTTAATTTTGTTTTGTTTTTTGAAATGGAGATCTCACTCTGTTGCCCAGGCTGGAGT 955
          || |||||| ||||||||| ||||||||||| ||||||| |||||||| ||||||||||||||
Db    898 GCCTATTTTTTAATTTTGTTTTGTTTTTTGAAATCTCACTCTTGACCTGGT 957

Qy    956 GCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTG 1015
          |||||||||||||||||||||||||||||||| |||||||||| ||||| ||||||||||
Db    958 GCAATGGGCCAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGCTCAAGCGATCTCCTG 1017

Qy   1016 TCTCAGCCTCCCAAGCAGCTGGGATTACGGGCGCCACCACCACCACGCCGCTAATTTTTG 1075
          ||||||||||||||||||||||||||| |||||||||||||||||||||| ||||||||||
Db   1018 TCTCAGCCTCCCAAGCACGTGGGATTACGGGCACCTG-CACCACCACCCCGCTAATTTTTG 1076

Qy   1076 TATTTTCATTAGAGGCGGGTTTCACCATATATTGGTCAGGCTGGTCTCAAAACTCCTGACCT 1135
          ||||| ||| ||||||||||||||||||||| |||||||||||| |||||||||||||||||
Db   1077 TATTTTCATTTAGAGGCGGGTTTCACCATATTGTCAGGCTGGTCTCGAACTCCTGACCT 1136

Qy   1136 CAGGTGACCCACCTGCCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCA 1195
          ||||||||| |||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   1137 CAGGTGACCCACCTGCCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCA 1194

Qy   1196 CCCAGCCGGCTAATTAGATAAAAAAAATATGTAGCAAATGTGGGGGGGGGTTCTTGCTATGTTGCCC 1255
          |||||||||||||||||||| |||||||| || |||||||||||||||||||||| ||||||||||
Db   1195 CCCAGCCGGCTAATTTAGATAATAAAAAATATGTAGCAAATGCTGGGGGATTACAGGCGTGA--CGCCTCA 1254

Qy   1256 AGGCTGGTCTCAAACTTCTGGCTTCAGTGCAATCTCCAAATGTCCAAATGAGCCACCACACCCAGCC 1315
          ||||||||||||||||| ||| ||||| |||||||| || |||||||||||||||||||||||||||
Db   1255 AGGCTGGTCTCAAACTTCTGGGTTCAGTGATTCTCATGATCTCCCTTCCAAATGCTGGGATTACAGGC 1314

Qy   1316 AGTCACATTTTTTAAACAGTTACATCTTTATTTTAGTATATACTGTAATACAACAATAAA 1375
          |||||| |||| |||||||||||||||||||||||||||||||| ||||| |||||| ||
Db   1315 AGTCACA-TTTTTAAACAGTTACATCTTTATTTACTGTAGAAAGTGATTACGATAAC 1373

Qy   1376 CATGTCAAACCTGCAAATTCAGTAGTAACAGAGTTCTTT 1414
          ||||| ||||| |||||||| ||||| ||||||||||||
Db   1374 ATGGCGGAACCTGCAAATTCGAGTAGTACAGAGTCTTTT 1412
```

RESULT 7
US-09-964-412-3

; Sequence 3, Application US/09964412
; Patent No. US20020129391A1
; GENERAL INFORMATION:
;   APPLICANT: de la Monte, Suzanne
;   Wands, Jack R.
;   TITLE OF INVENTION: Transgenic Animals and Cell Lines for
;   Screening Drugs Effective for the Treatment or Preventio
;   of Alzheimer's Disease
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox, P.L.L.C.
;   STREET: 1100 New York Ave., Suite 600
;   CITY: Washington
;   STATE: DC
;   COUNTRY: USA
;   ZIP: 20005-3934
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC Compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/964,412
;   FILING DATE: 28-Sep-2001
;   CLASSIFICATION: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Esmond, Robert W.
;   REGISTRATION NUMBER: 32,893
;   REFERENCE/DOCKET NUMBER: 0609.4370000
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 202-371-2600
;   TELEFAX: 202-371-2540
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1381 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: both
;   MOLECULE TYPE: cDNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-964-412-3

Query Match   74.9%;   Score 1080.2; DB 10; Length 1381;
Best Local Similarity   94.4%;   Pred. No. 3e-217;
Matches 1317; Conservative   0; Mismatches 53; Indels 25; Gaps 18;

```
Qy    2 TTTTTTTTTTGAGATGGAGT...                                           61
        ||||||||||| ||||||||
Db    1 TTTTTTTTTTTGAGATGGAGAGTTTTCGCTCTGTTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT 60

Qy   62 CTCAGCTCACCGGCAACCTCCGCGTTCAGGCGATTCTCTGCCC                         121
        |||||||||||||||||||| |||||| |||||||||||||||
Db   61 CTCAGCTCACCGGCAACCTCGCGCGTTCCGGCGATTCTCTGCCC                        120

Qy  122 AGTAGCTGGGATTACAGGCATGTGACCACCACGGTTCAGGCTAATTTTTGTATTTTTTTAGT     181
        |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db  121 AGTAGCTGGGATTACAGGCATGCA-CCACGCTCGGCCTAATTTTTGTATTTTTTTAGT          179

Qy  182 AGAGATGGAGGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT      239
        |||||||||     ||||||||||||||||||||||||||||||||||||||||||
Db  180 AGAGATGGAGTTTAACTCAGGGTTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT          239

Qy  240 CCCTCCGTCTCGGCCGTCCCAAAGTGCT---AGATACAGGACTGGCCACCATGCCCGG-CT      295
        ||||| |||||||||||||||||||||   ||||||||||||||||||||||||||| ||
Db  240 CTCCGGTCTCGGCCGTCCCAAAGTGCTGAGATTACAGGCGTCACCATGCCCGGCT             299

Qy  296 CTGCCTGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAGCCTGGTCTC      354
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  300 CTGCCTGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAGCCTGGTCTC       359

Qy  355 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAG      414
        ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db  360 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG       418

Qy  415 CCGTGGCCTGGGCCTTTATTTTTTTATTTTTTAAGACACAGGGTGTCCCACTCTTTACCCAGG   474
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  419 CCGTGGCCTGGGCCTTTATTTTTTTATTTTTTAAGACACAGGGTGTCCCACTCTTTACCCAGG    478

Qy  475 ATGAAGTCCAGTGGTGTGATCACAGCTCACTGCAGCCTCCGACCTCCTGAGATCAAGC-AT     533
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  479 ATGAAGTCCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCGAAT     538

Qy  534 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGACCACTACACCTGCCTA       593
        ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  539 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGACCACTACACCTGG-TA      597

Qy  594 ATTTTTTTATTTTTATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGA          653
        |||||| ||||||||||||||||||||||||||||||||| ||||||||||||||
Db  598 ATTTTTTATTTTTATTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA           656

Qy  654 GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC      713
        ||||||||||||||||||||||||||||| ||||| ||||||||||||||| ||||||
Db  657 GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC      716

Qy  714 TGCCCCACCTCCTGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCCGGCTAATTTTTT     773
        |||||| ||||||||||||||||| |||||||||||||||||||||||| |||||||
Db  717 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACTAATTTTTT           776

Qy  774 GTATTTTTAGTAGAGATGGGG--TTCACCATGTTGGCCAGGTTGGTCTCGATCTCTGGACC    832
        ||||||||||||||||||||  |||||||||||||||||||||||||||| |||||||||
Db  777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGTTGATCTCGATCTCTTGGACC      836

Qy  833 TTGTGATTCTGCCTGCCTCGGCCT-CCCAAGTGCTGGGATTACAGGCGTGAGCCACCACG     891
        |||||| |||||||||| ||| |  |||||||||||||||| |||||||||||||||||
Db  837 TTGTGAATCTCGCCTGCCTCGGCC--GTCGTGACTCCAC                          894

Qy  892 CCCGGCTTATTTTTTAATTTTTTGTTTGTTTGGAAATGGAATCTCACTCTGTTACCAGGCTG     951
        ||| ||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  895 GCCGGCCTATTTTTAATTTTTTGTTTGTTTGAAATGGAATCTCACTCTGTTACCAGGCTG      954

Qy  952 GAGTGCAATGGCACAAATCTCGGCTCACTGCCAACCTCTGCCTCCCGGGCTCAAGCGATTCT    1011
        ||||||||||| |||||||||| ||||||||||||||||||||||| |||||
Db  955 GAGTGCAATGG-CAAAATCTCGGCTACTCGGCGAACTCTGCCTCCTCGGGG-TCAAGCGATTCT  1012

Qy 1012 CCTGTCTCAGCCTCCCAAGCAGCTGGGACCACACCACCCCGCCCGGCTAATT            1071
        |||||||||||||||||||||| || ||||||||||||||||||||||||
Db 1013 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGG--ACCTGCCACACACCCCGGCTAATT    1070

Qy 1072 TTTGTATTTTCATTAGAGGCGGGGTTTCACCATATATTGTGTCAGGCT-GGTCTCAAACTCCT   1130
        ||| |||||||||| |||||||||  |||||||||||||||||||| |||||||||
Db 1071 TTTGTATTTTCATTAGAGGGCGGG--TTTACCATATTTGTGTCAGGCTGGGTCTCAAACTCCT  1128

Qy 1131 GACCTCAGGTGACCCACCTCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA   1190
        |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db 1129 GACCTCAGGTGACCACCTCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA   1188

Qy 1191 CCTCACCCAGCCGGCTAATTTTTAGATAAAAAAAATAGTAGCAATGGGGTTCTTGGTATGT   1250
        |||||||||||||||||||||| |||||||||||||||||||| |||| ||||||||||||
Db 1189 CCTCACCCAGCCGGCTAATTTTTGGATAAAATATGTAGCAATGGGG---TCTGCTATGT    1246

Qy 1251 TGCCCAGGCTGGTCTCGAAACTTCTGGCTCAGTGCAATCCTTCCAAATGAGCCACCACACC   1310
        |||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||||
Db 1247 TGCCCAGGCTGGTCTCAAACTTCTGGCTCAGTCAATCCTTCCAAATGACCACCACACC   1306

Qy 1311 CAGCCAGTGCACATTTTTTAAACAGTTTACATCTTTATTTTAGTATACTAGAAAGTAATACA   1370
        |||||||| |||| ||||||||| ||||||||||| ||||||||||||||||||||||||||
Db 1307 CAGCCAGTCACATTTTTTAAACAGTTACATCTTTTATTTTAGTATACTAGAAAGTAATACA   1366

Qy 1371 ATAAACATGTCAAAC   1385
        |||||||||||||||
Db 1367 ATAAACATGTCAAAC   1381
```

```
Qy     62  CTCAGCTCACCGCAAACTTCCGCCTCGCCTCAGCCTCCCC                                                121
           ||||||||||||||||||||||||||||||||||||||||
Db     61  CTCAGCTCACCGCAAACTTCCGCCTCGCCTCAGCCTCCC                                                 120

Qy    122  AGTAGCTGGGATTACAGGCATGTGCACCACCGCTCGGCTAATTTTGTATTTTTTTTTTAGT                           181
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AGTAGCTGGGATTACAGGCATGTGCA-CCACGCTCGGCTAATTTTGTATTTTTTTTTAGT                            179

Qy    182  AGAGATGGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGAT                           239
           |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    180  AGAGATGGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT                           239

Qy    240  CCCTCCGTCTCGGCCTCCCAAAGTGCT---AGATACAGGACTGGCCACCATGCCCGG-CT                             295
           ||||||||||||||||||||||||||    ||||||||||||||||||||||||||| ||
Db    240  CTCCCGTCTCGGCCTGCCCAAAGTGCTGTGAGATACAGGCGCATGAGCCACCGCAGATGAT                           299

Qy    296  CTGCCTGGCTAATTTTTGTGGTTAGAAAACAGGGTTTCACTGATG-TGCCCAAGCTGGTCTC                          354
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    300  CTGCCCGGCTAATTTTTGTGGTTAGAAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTC                          359

Qy    355  CTGAGCTCAAGCGATCTCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCGAG                          414
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360  CTGAGCTCAAGCGATCTCCCCCGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG                                 418

Qy    415  CCGTTGCCTGGCCTTTTATTTATTTTATTTATTTTATTTTTAAGACACAGGGTCTCACTCTTACCCAGG                   474
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419  CCGTTGCCTGGCCTTTTATTTATTTTATTTATTTTATTTTTAAGACACAGGGTCTCACTCTTACCCAGG                   478

Qy    475  ATGAAGTGCAGTGGTGTGATCACAGCTCACTCGCAGCCTTCAACTCCTGAGATCAAGC-AT                           533
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    479  ATGAAGTGCAGTGGTGTGATCACAGCTCACTCCTGAGATCAAGCGCAAT                                       538

Qy    534  CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACCTGAGCTGGCTA                         593
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    539  CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACCTGG-TA                              597

Qy    594  ATTTTTATTTTATTTTTATTTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGA                     653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598  ATTTTTATTTTATTTTTAATTTTTTGAGACAGAGTCTC-ACTCCTGTCACCCAGGCTGGA                            656

Qy    654  GTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC                            713
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657  GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCCGGGGTTCAAGTTATTCTCC                                  716

Qy    714  GTATTTTTTAGTAGAGATGGGG-TTCACCATGTTGCCCAGGCTGGTCTCGAACTCCTGACCTCAGGGT                    773
           |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    717  TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCTAGTTAATTTTTTTT                       776

Qy    774  GTATTTTAGTAGAGATGGGG-TTCACCATGTTGTCCCAGGCTGGTCTCGGACCTCAGGGT                            832
           ||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    777  GTATTTTTTAGTAGAGAATGGGGTTTCACCATGTTGATCCTCGGACC                                        836

Qy    833  TTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG                            891
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    837  TTGTTGATCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAG--GTCGTGACTCCAC                               894

Qy    892  CCCGGCCTTATTTTTTATTTTTAATTTTTTGTTTTGAAATGGAAATCACTCTGTTCCAGGCTG                         951
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    895  GCCGGCCTTATTTTTTAATTTTTGTTTTGTTTTGTTTTGAGATCACTCTGTTTACCCAGGTCG                         954

Qy    952  GAGTGCAATGGCCAAAATCTCGGCTCACTGCAACTCTCTGCCTCCCGGGGTTCAAGCGGATTCT                        1011
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    955  GAGTGCAATGG-CAAAATCTCGGCTACTGCCGCACCTTCGCCTCCCGGGGG-TCAAGCGATTCT                        1012

Qy   1012  CCTGTCTCAGCCTCCCAAGCAGTCTGGGATTACAGGCACCCACCACCACGCCCGGCTAATT                          1071
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1013  CCTGTCTCAGCCTCCCAAGCAGTGGGGATTACGGG--ACCTGACCACACCCCGGCTAATT                           1070

Qy   1072  TTTGTATTTTTTCATTAGAGACGGGG--TTTACCATATTGTTCAGGCTGTCTCAAAACTCCT                         1130
           ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db   1071  TTTGTATTTTTCATTAGAGGCGGGGGGTTTACCATATTGTTCAGGCTGGTCTCAAAACTCCT                         1128

Qy   1131  GACCTCAGGTGACCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA                            1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1131  GACCTCAGGTGATTACAGGCGTG--TTTACCATATTTGTTCGACAGCTTCAGCGAGATGATCGC                       1190
```

```
Query Match          40.8%;  Score  587.8;  DB  9;  Length 65608;
Best Local Similarity 70.6%;  Pred. No. 4.9e-114;
Matches   940;  Conservative   0;  Mismatches 347;  Indels  44;  Gaps  10;
```

```
Qy      4  TTTTTTTTGAGATGGGAGAGTTTCGCCTCTTTGTTCCCCCAGGCTGGAGTGCAATGGCAAATCT                         63
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39993  TTTTGTTTTGAGATGGGAGAGTTTT-GCTCTGTTGCCCAGGCTGAAGTGTAATGGCACGGATCT                        39935

Qy     64  CAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG                             123
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39934  CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAGCAATTCTCCTGACCTCAGCCTCCCGAG                           39875

Qy    124  TAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG                               183
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39874  TAGCTGGGATTACAGGCATGTGCCACCACCACCTGGCTAATTTTTGTATTTTT---AGT                             39819

Qy    184  AGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCGC                           243
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39818  AGACGGGGTTTCTCCATGTTGGTCAGGCTTGGTCTCGAACCCCGACCTCAGATGATCCGC                            39759
```

```
Qy      244 CCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCTCGGCTCCTGG 303
            ||  || |||||| |||||||  ||  |  |||| | |   | || || ||| | |||||||
Db    39758 CCACCTTGGCTCCCAAAGTGCT-GGGATTACAGGTGTGAGTTTTACAAGAAAAACATTTA 39700

Qy      304 CTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGCTCA 363
            ||                    || ||||| |||       | |  | || | ||||||
Db    39699 AGTATGAAAGCAGTACCTCCGGAATATGCACGGCCCCTCCAAACCCAGCCTTGAAGACTC 39640

Qy      364 AGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTA------------------ 403
            ||                   | ||  | ||  ||||  |
Db    39639 TTCACTTGTCTCTGGCCTGGGCGCCTTCCAGCCTCCCAAAGTGCTGGGATTCTCTT 39580

Qy      404 CAGGCGTGCAGCGCCGTGGCCTGGCTTTTTTTTTTTTTATTTTTTTTTTTAAGACAGG 463
            |||| |||| | | | || ||  |||||||||||||| |||||||||||
Db    39579 CCCCCTCCCATCCCTGCCTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGTCTTGTTT 39520

Qy      464 TCTTACCCAGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGA 523
            |||| |||||||||| | | || || ||||| ||||| |||||| ||| ||||||||||
Db    39519 TGTTGCCCAGGCTGGAGTGCAGTGGTGCAAACATGCCATCACAGCCTTGATCTCTTGG 39460

Qy      524 GATC-AAGCATCCTCCGTCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACCACT 582
            ||||  |  |  | |||   ||||||||||||| ||||| | || | | |  |||  |||
Db    39459 GCTCAAAGGGATCCTCCCACCTCAGCCTCCCGAGTAGCTGGGACTATAGGTGCATGCCACC 39400

Qy      583 ACACCTGGCTAATTTTTTATTTTTTAGTAGAGACAGGGTCTCAACTCTGTCA 642
            |||||| ||||||||||||   ||||||  | | | | |||| | |||| |
Db    39399 ATGCCCAACTAATTTTTTTT-----TTTTTGAGATGGAGTTTCGCTCTTGTTG 39348

Qy      643 CCCAGGCTGGAGTGCAGTGGCGCCAATCTTGGCTCACTGCAACTCTGCCTCCCGGGTTCA 702
            |||||||||| ||||| |  | || | |||||||||||||| | ||||||| |||||||
Db    39347 CCCAGGCTGGGCTGCAATGGTGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCA 39288

Qy      703 AGTTATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCCTAG 762
            || || | |  ||||||||||||  |||| |||| ||||||||| |  |||||| |||
Db    39287 AGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGCCCACCACACC 39228

Qy      763 CTAATTTTTTGTATTTTTAGTAGAGATGGGG-TTCACCATGTTGGTCAGGCTGGTCTTG 821
            |||||| |||| |||| |||  |||||||||  || || | || | |||||| || ||
Db    39227 TGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGA 39168

Qy      822 ATCTCTGGACCT--TGTGATCCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 879
            ||||||  | |   ||| |  | |||  ||||  ||||||||||| || |||||| |||
Db    39167 AACTCCTGACCTCAGGATCTCCCTGCCTCGGCCTCCCAAAGCAATTCTCCTACCTCAGCCTCCCGAG 39108

Qy      880 TGAGCCACCACGCGCCCGGCCTT--AtTTTTTAATTTTTGTTTGTTGTGTTTGAATTCACTC 937
            |||| |||||| |   |||||    | ||||||| ||||||| || ||   || |||  ||
Db    39107 TGAGCCACTGTGGCTGGCCAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGC 39048

Qy      938 TGTTTACCCAGGTGGAGTGCAAATCGGGTCTCACTGCAACCTCTGCCTCCCG 997
            |||||||| |  || |||| |||  |||||||||| | || |||| ||| |
Db    39047 TCTTTCCCAGGCTGGAGTGCAGTGGCACTTCGGCTCACTGCAACCTCCGCCTCCCGG 38988

Qy      998 GCTCAAGCGGATTCTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCACCTGCCACC 1057
            |||||| ||| | ||| | ||||||||||||| ||||| ||||| ||||| |||| ||| |||
Db    38987 GTTCAAGCGGATTCTCCTCAGTGGCATAGCAGCTGGGACTACAGGCGTGGTCTCCTGAGCTCA 38928

Qy     1058 ACACCCGCTAATTTTTGTATTTTTCATTGAGGGTTCACCATATTGTGCAGGCTG 1117
            ||| || |||||||| |||||| |   |  |||  | |||| ||| | |||||
Db    38927 ACGCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTG 38868

Qy     1118 GTCTCAAAACTCCTGACCTCAGGTGATCCGACCCACGCCTCCTCAGCCTCGGGATTA 1177
            |||||  |||||||||| ||| |||||| | | | || |||| | |||| |||||
Db    38867 GTCTTGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 38808

Qy     1178 CAGGCGTGAGCCACCACCTCACCCAGCC----GGCTAATTTAGATAAAAAAATATGTAGCAAT 1233
            ||||| || |||||||| |||| |||||     |||| || |||| |||||  | |   ||
Db    38807 CAGGCATGAGCCACCACCGCCCTAATTAATAAAAAATTTTTTGTAGAG 38748

Qy     1234 GGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAACTCCTGGGCTCAAGCGATCAATCCTCC 1293
            ||||| | || |  || || ||| ||  || |  | ||| |||||||||| | |  ||||||
Db    38747 ACAGGATTCACTATGTTGTCCAGACTGGTCTCGAACTCCTGGGCTTAAGGGATGCTGTCCT 38688

Qy     1294 AAATGAGCCAC 1304
             || ||||| |
Db    38687 GTCTCAGCCTC 38677
```

```
Query Match      40.8%;  Score 587.8;  DB 10;  Length 65608;
Best Local Similarity   70.6%;  Pred. No. 4.9e-114;
Matches  940;  Conservative  0;  Mismatches 347;  Indels 44;  Gaps 10;

Qy        4 TTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTCCCAGGCTGGAGTGCAATGGCGCAATCT 63
            ||||||||| ||||||||||||  |||||||||| ||||| |||| |  | || ||| ||||
Db    39993 TTTTTTTTTGAGATGGAGTTTT-GCTCTTGTTGCCCCAGGCTGAAGTGCAGTGGCACGATCT 39935

Qy       64 CAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAG 123
            |||||||| ||||| ||| ||||||||| ||||| | ||| | ||||| |||||||| ||
Db    39934 CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAAGCAATTCTCCTACCTCAGCCTCCCGAG 39875

Qy      124 TAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTTGTATTTTTTAGTAG 183
            ||||||| ||||||||| || || |||||  | |||||||||| |||| ||||| ||||
Db    39874 TAGCTGGGATTACAGGCGTGCTCACCACGCCTAATTTTTGTATTTTT--AGT 39819

Qy      184 AGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCGCCT 243
            ||| ||||| ||  ||||||| ||||| ||||||| ||||| ||| |  |||| ||| |||
Db    39818 AGACGGGGTTTCTCCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAGATGATCCGCC 39759

Qy      244 CCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCTCGGCTCCTGG 303
            ||  || |||||| |||||||  ||  |  |||| | |   | || || ||| | |||||||
Db    39758 CCACCTTGGCTCCCAAAGTGCT-GGGATTACAGGTGTGAGTTTTACAAGAAAAACATTTA 39700

Qy      304 CTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGCTCA 363
            ||                    || ||||| |||       | |  | || | ||||||
Db    39699 AGTATGAAAGCAGTACCTCCGGAATATGCACGGCCCCTCCAAACCCAGCCTTGAAGACTC 39640

Qy      364 AGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTA------------------ 403
            ||                   | ||  | ||  ||||  |
Db    39639 TTCACTTGTCTCTGGCCTGGGCGCCTTCCAGCCTCCCAAAGTGCTGGGATTCTCTT 39580

Qy      404 CAGGCGTGCAGCGCCGTGGCCTGGCTTTTTTTTTTTTTATTTTTTTTTTTAAGACAGG 463
            |||| |||| | | | || ||  |||||||||||||| |||||||||||
Db    39579 CCCCCTCCCATCCCTGCCTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGTCTTGTTT 39520

Qy      464 TCTTACCCAGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGA 523
            |||| |||||||||| | | || || ||||| ||||| |||||| ||| ||||||||||
Db    39519 TGTTGCCCAGGCTGGAGTGCAGTGGTGCAAACATGCCATCACAGCCTTGATCTCTTGG 39460

Qy      524 GATC-AAGCATCCTCCGTCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACCACT 582
            ||||  |  |  | |||   ||||||||||||| ||||| | || | | |  |||  |||
Db    39459 GCTCAAAGGGATCCTCCCACCTCAGCCTCCCGAGTAGCTGGGACTATAGGTGCATGCCACC 39400
```

```
RESULT 10
US-09-962-832-119/c
Sequence 119, Application US/0996283
Patent No. US2002011082A1
;  GENERAL INFORMATION:
;  APPLICANT: Ebner, Reinhard
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;  FILE REFERENCE: 689299-74
;  CURRENT APPLICATION NUMBER: US/09/962,832
;  CURRENT FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,077
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,280
;  PRIOR FILING DATE: 2000-09-25
;  NUMBER OF SEQ ID NOS: 259
;  SOFTWARE: Patentin version 3.0
;
;  SEQ ID NO 119
;  LENGTH: 65608
```

```
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  OTHER INFORMATION: n=a,t,g or c
US-09-962-832-119

Query Match    40.8%;  Score 587.8;  DB 10;  Length 65608;
Best Local Similarity  70.8%;  Pred. No. 4.9e-114;
Matches 940;  Conservative 0;  Mismatches 347;  Indels 44;  Gaps 10;
```

```
Qy      4 TTTTTTTTTGAGATGGAGTTTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAATCT    63
Db  39993 |||| ||||||  |||||| |||||||||||||  |||  ||||| |||||||||||||||||| 39935

Qy     64 CAGCTCACCGCCAACCTCCGCCTCCCGGGTTCAAGCGGATTCTCCCGCCTCAGCCTCCCCAG   123
Db  39934 ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||| 39875

Qy    124 TAGCTGGGATTACAGGCATGCACCACCACGCCCGGCTAATTTTTGTATTTTTTTTAGTAG    183
Db  39874 |||||||||||||||||| ||||||||||||||||||| |||||||||||||||||||||| 39819

Qy    244 CCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGCCACCACGCCCGGCCTGCCTGG    303
Db  39758 ||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||| 39700

Qy    304 CTAATTTTTGTGGTAGAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGCTCA    363
Db  39699 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 39640

Qy    364 AGCAGTCCACTGCCTCAGCCTCCGGAATATGCACGGCGCCCTCCAACCCAGCCTGTGGGGATTA    403
Db  39639 ||||||||||||||||||||||||||||||||||||||||||||||          ||||| 39580

Qy    404 CAGGCGTGAGCCACCGCCGCCTGGCCTTTTTT     463
Db  39579 ||||||||||||||||||||||||||||||||| 39520

Qy    464 TCTTACCCAGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCTGA    523
Db  39519 ||||||||||||||||||||||||||||||||||||||||||||||||||||||| 39460

Qy    524 GATC-AAGCAACCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACCACACT    582
Db  39459 ||||  |||||||||||||||||||||||||||||||||||||||||||| 39400

Qy    583 ACACCTGGCTAATTTTTATTTTTATTTTTTTGAGACAGAGTCTCAACTCTGTCA    642
Db  39399 ||||||||||||||||||||||||||||||||||||||||||||||||||||| 39348

Qy    643 CCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCA    702
Db  39347 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 39288

Qy    703 AGTTATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCCCACCACGCCCTAG    762
Db  39287 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 39228

Qy    763 CTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTTGATCTTG    821
Db  39227 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 39168

Qy    822 ATCTCTGGACCTT--TGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG    879
Db  39167 ||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||| 39108

Qy    880 TGAGCCACCACGCCCGGCCTT--ATTTTTAATTTTTGTTTGTTTGAAATGGGAATCTCACTC    937
Db  39107 ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||| 39048

Qy    938 TGTTACCCAGGCTGGAGTGCAAATGGCGCAAATCTCGGCTCACTGCAACCTCTCGCCTCCGG    997
Db  39047 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 38988

Qy    998 GCTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACC   1057
Db  38987 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 38928

Qy   1058 ACACCCGGCTAATTTTTGTATTTTTCATTAGAGACGGGGTTTCACCATGTTGTCAGGCTG   1117
Db  38927 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 38868

Qy   1118 GTCTCAAACTCCTGACCTCAGGTGACCCACCTCGCCTTCCAAAGTGCTGGGATTA   1177
Db  38867 |||||||||||||||||||||||||||||||||||||||||||||||||||||||| 38808

Qy   1178 CAGGCGTGAGCCACCGCCGCCCAGCC----GGCTAATTTAGATAAAAAAAATATGTAGCAAT   1233
Db  38807 ||||||||||||||||||||||||    |||||||||||||||||||||||||||||||| 38748

Qy   1234 GGGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAACTTCTGGCTTCAATCCTTCC   1293
Db  38747 ACAGGATCACTATGTTGTCCAGACTGGTCTTGAACTCCTGGGCTTAAGGGATCGTCCT 38688

Qy   1294 AAATGAGCCAC   1304
Db  38687 GTCTCAGCCTC 38677
```

```
RESULT 11
US-09-954-456-2257
Sequence 2257, Application US/0995456
Patent No. US20020115057A1
GENERAL INFORMATION:
  APPLICANT: Young, Paul
  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
  TITLE OF INVENTION: Sets
  FILE REFERENCE: 689290-76
  CURRENT APPLICATION NUMBER: US/09/954,456
  CURRENT FILING DATE: 2001-09-18
  PRIOR APPLICATION NUMBER: US/60/233,617
  PRIOR FILING DATE: 2000-09-18
  PRIOR APPLICATION NUMBER: US/60/234,052
  PRIOR FILING DATE: 2000-09-20
  PRIOR APPLICATION NUMBER: US/60/234,923
  PRIOR FILING DATE: 2000-09-25
  PRIOR APPLICATION NUMBER: US/60/235,134
  PRIOR FILING DATE: 2000-09-25
  PRIOR APPLICATION NUMBER: US/60/235,637
  PRIOR FILING DATE: 2000-09-26
  PRIOR APPLICATION NUMBER: US/60/235,638
  PRIOR FILING DATE: 2000-09-26
  PRIOR APPLICATION NUMBER: US/60/235,711
  PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: US/60/235,720
  PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: US/60/235,840
  PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: US/60/235,863
  PRIOR FILING DATE: 2000-09-27
  NUMBER OF SEQ ID NOS: 2276
  SOFTWARE: Patentin version 3.0
SEQ ID NO 2257
  LENGTH: 62944
  TYPE: DNA
  ORGANISM: Homo sapiens
US-09-954-456-2257

Query Match      36.4%;    Score 524.4;  DB 10;  Length 62944;
Best Local Similarity  71.9%;   Pred. No. 8.6e-101;
Matches  963;  Conservative  0;  Mismatches  321;  Indels  56;  Gaps  19;
```

```
Db 39107  TGAGCCACTGTGCCTGGCCTGGCCCAATTTTTTTTTTTTTTTTTTTTTGAGACAAAGTCTCGC 39048
Qy   938  TGTTACCCAGGCTGGAGTGCAAATGGCGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGG   997
          |||   ||||||||| ||||||| |  |    | || ||||||||| |||||| |||||| |||
Db 39047  TCTTTCCCAGGCTGGAGTGGAATGGCACCACCTCACTGCAACCTCTGCCTTTCGG 38988
Qy   998  GCTCAAGCGGATTCTCGTGTTCACGCCTTCCCCAAGCACAGCCACCTGCCACC 1057
Db 38987  GTTCAAGCGATTATTCCTGCCTCAGTCTCCCAAGTAGCTAGGACTACAGGCGCACC 38928
Qy  1058  ACACCCCCGCGTAATTTTTTGTATTTTTAGTAGAGACGGGGGTTTCACCATGTTGGTCAGGCTG 1117
Db 38927  ACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACAGGGGTTTCACCATGTTGGCCAGGCTG 38868
Qy  1118  GTCTCAAACTCCTGACCTCAGGTGACCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 1177
Db 38867  GTCTTGAACTCCTCAACTTCAGGTGATCACCACCGCGTCGGCCTCCCAAAGTGCTGGGATTA 38808
Qy  1178  CAGGCGTGAGCCACCTCACCCAGCC---GGCTAATTTAGTAGAAATAATAATTGTAGCAAT 1233
Db 38807  CAGGCAATGAGCCACCACCACCTGGCCCTCTGCCCAACTAATTAAAAAAAAATTTTTGTAGAG 38748
Qy  1234  GGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAACTTCTGGCTTCATGCAAATCCTTCC 1293
Db 38747  ACAGGANTCTCACTATGTTGTCCAGGACTGGTCTTGAACTCCTGGGCTCAAGCGATCGTCCT 38688
Qy  1294  AAAATGAGCCAC 1304
Db 38687  GTCTCAGCCTC 38677
```

```
Qy     1  TTTTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAAATGGCGCAA   60
Db 59357  TATTTTATTTTTGAGATGGAGTTTT-GCTCTTCTCGGCCCAGGCTGGAGTGCGCAA 59415
Qy    61  TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC  120
Db 59416  TCTCGTCTCACTGCAACCTCCGCCTCCCGAGTTCAAATAGATTCTCCTGCCTCAGCCT--- 59472
Qy   121  CACGTAGCTGGGATTACAGGCGCACCGCCACCACGCCCGGCTAATTTTTTGTATTTTTTAG  180
Db 59473  TAGTAGGTGTGATTACAGGCGTGAGCCACCGCGCCCAGCCTAATTTT---TGTATTTTTAC 59529
Qy   181  TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAAACTCCTGACCTCAGGTGATC  240
Db 59530  TAGAGACAGGGTTTCACCATGTTGGCCAGGCTGTCTTGAACTCCTGACCTCAAGTGATC 59589
Qy   241  CCTCCGTCTCGGCCTCCCAAAGTGCTAG----ATACAGGACTGGGCACCAGGCCCGGCTC  296
Db 59590  TGCCTACCCTCGGCCTCCCAAAGTGCTGGGATTACACACAGGGCACGCAGCCCTC 59649
Qy   297  TGCCTGGCTAAT-TTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCC  355
Db 59650  TTTTTTCTTTTTCTTTTTCTTTTTTCTGAGGCAGGTCTCGCTCTGTCACCCAGGCTGGAGTGC 59709
Qy   356  TG----AGCTCAAGCAGTCCACCTGTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTG  411
Db 59710  AGCAGCAGGATCAATAGCTCACTGAGCTTCGATCTCCCGGGTTCAAGTGATCCTCCCAGCT 59769
Qy   412  CAGGCCGTGCCTGGCCTTTTTTTTATTTTTATTTTTTTTTTTTTTTTGAGATGGA 471
Db 59770  AATTTTTTTTTTTTTATTTTTATTTTTTTTTTTTTTTTTTGAGATGGA-GTCTGGCTCTGTCGCCC 59828
Qy   472  AGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCAACTCCTGAGATCAAGC  531
Db 59829  AGACTGGAGTGCAGTGGCGCACAATCTCAGCTCACTGCAACCTT-GCCTCCTGGATTCAAGC 59887
Qy   532  ATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTAAAAGACATGCACCACTACACCTGGC  591
Db 59888  AATTCTCTGCCTCAGCCTCTGA-GTAGCTAGGATTACAGGCGCCACCACCGCCCGGA 59946
Qy   592  TAATTTTTTA--TTTTTTATTTTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGC  649
Db 59947  TAAAATTTTGGGTTTTTTTTTTTTTTTTTTTTTTTTTAGACAGGGTCTC-ACCCTGTCGCCCAGGC 60005
Qy   650  TGGAGTGCAGTGGCGCCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATT  709
Db 60006  TGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAAGCTCTGCCTCCCGGGTTCACGGCCATT 60065
Qy   710  CTCCTGCCC-------CAGCCTCGTGAGTAGCTGGGACTACAGGCGCC-----CCACACGCC  759
Db 60066  CTCCCGCCTCCCCAACCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCACCACCATGC 60125
Qy   760  TAGCTAAT-------TTTTGTATAATTTTTTAGTAGAGACGGGGGTTTCACCGTGTTGGCCAGGATG  815
Db 60126  CAGTTAATTTTTTTGTATTTTTTAGTAGAGACGGGTTTCACCGTGTTAGCCAGGATG 60185
Qy   816  ATCTTGATCTCTGGAACCTGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTAGGATTACA  875
Db 60186  GTCTCGAAATCTCCGACCTGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACA 60245
Qy   876  GGCGTGAGCCACCGCACCGCCCGCCGCTTTATT---------TTTAATTTTTTGTTGTTGAAAT  925
Db 60246  GGCGTGAGCCACCGCACCGCCCGCCTGTTTTTCGTTTGCCATTTTTTTTTTTTTGTTTTTTTGGACG 60305
Qy   926  GGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAATGGCCAAAATCTCGGCTCACTCAACC  985
Db 60306  GGCCTTGAGCGGCCACCGCCCTGCGTTTATTGGAGTTGCCAATGGTCACTCAGCTCACTACAAAC 60365
Qy   986  TCTGCCTCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGCTGGAGTAGCTGGGATTACCG 1045
Db 60366  TCTGCCTCCCAGGTTCAAGCGGATTCTCCTGCCTCAGCCTCCCAAGGGAGTAGCTCACTGCAAAC 60425
```

```
Qy     1046 GCACTGCCACCACCCCCCTAA-TTTTGTATTTTCATTAGAGGGGGGGGTTTCACCAT 1104
            |||          |||      |  ||  ||||||  | |||||||||||||||||
Db    60426 GCCGTGCCACCACCATGCCCGGCGGCTAATTTTTGTATTTTTAGTAGAAATGGGGTTTCACCAT 60485

Qy     1105 ATTTGTCAGCGCTGGTTCTCAAACTCCTGACCTCCTCAGGTCACCGCCTTCCAA 1164
            ||||||| |||||  |  ||||||||||   |||| ||||||||||||||||||
Db    60486 ATTTGGTCAGCGCTGGTTCTCAAACTCTTGACCTC--GTGATCACCGCCTTCCAA 60543

Qy     1165 AGTGGCTTGGGGATTACAGGCGTGAGCCAACCTCACCCAGCCGGCTAATTTAGATAAAAATA 1224
            ||||||||||||||||||||||||||||||||||||| ||| ||||||||     |
Db    60544 AGTGCTGTGGGATTACAGGCATAAACCACTGCGCCCGGCTG----------------T 60596

Qy     1225 TGTACCAAATGGGGGTCTTGCTATGTTGGCCAAGGCTGCTGTCATGC 1284
            |||||||||||||||||| |||||  ||||||||||||||| |
Db    60597 TTAATAGAGATGAGGTCACACTGTTGTTGCCCAGGCAGACAGTCTCGGAACTTCCTGAATTCACAC 60656

Qy     1285 AATCCTTCCAAATGAGCCAC 1304
            |||||||||||||||
Db    60657 AATCTGCCTGCCTCAGCCTC 60676
```

RESULT 12
US-09-764-847-1157
Sequence 1157, Application US/09764847
Patent No. US20020132767A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC009
CURRENT APPLICATION NUMBER: US/09/764,847
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 2003
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 1157
LENGTH: 21470
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-847-1157

Query Match   35.6%; Score 513; DB 10; Length 21470;
Best Local Similarity 71.2%; Pred. No. 1.7e-98;
Matches 885; Conservative 0; Mismatches 290; Indels 68; Gaps 13;

```
Qy        4 TTTTTTTTGAGATGGAGTTTCTCCATTGTTGGGTCAGGCTGTCGGAACTCCCGACCTCAGTGATCCTC 63
            ||||  |  ||||||||||||| ||||||||||||  |||| |||||  ||||||  ||
Db     2359 TTAATTTGAGATGGCAGTTTT-GCTCTTGTTGTCCCTAGGCTGGAGTGCAGTGGCACGAATCT 2417

Qy       64 CAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG 123
            |||     |||||||||  ||||||||||||||||  ||||||||||||||||||||||
Db     2418 TGGCTCGTTGCAACCTCTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAG 2477

Qy      124 TAGCTGGGATTACAGGCGTGCACCCACGCTCGGCCGGCTAATTTTGTATTTTTTAGTAG 183
            ||||||||||||||||   ||||  |||| |||  |||||||||||| |||||  |||
Db     2478 TAGCTGGGATTATAGGCACGTGCCACCACTGCCTGGCTAATTTT---TGTATTTTTAATTAG 2534

Qy      184 AGATGGAGTTTCTCCATTGTTGGTCAGGCTGTCTCCGAACTCCCGACCTCAGATGATCCT 243
            |||||||||||| || |||||||  ||||||||| |  ||||||  |||||| |||||
Db     2535 AGACGGGGTTTCGCCATAATTGGCCAGGCTGGTCTCGAACTCCTGACCTCGTGATCCGC 2594

Qy      244 CCGTCTCCGGCCTCCCAAAGTGCT---------AGATACAGGACTGGCCACCATGCCCGGC 294
            ||| ||||| |||||||||||  |         |||  ||| |||  |||||||||||||
Db     2595 CCACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCTACCAGCGCCCGGAAC 2654

Qy      295 TCTGCCTGGCTAATTTTTGTGTAGAAAACAGGGTTTCACTGATGT---------- 339
            |||||| ||||||||| |  ||||||||||||||||||||||
Db     2655 TTATAGAATTTAACCATAATTATTTATCACAGAATTACAGATTGAAATCGAAT 2714

Qy      340 ----GCCCAAGCTGGTCTCGTCTCTGAGCTCAAGCGAGTCCACCTGCCTCCCAAAGTGC 395
                ||||  |||||   |||| |||||| | |||||||| |||| | | ||||||||
Db     2715 TCTCCTTCCCTCTTTAACTGCTGGGCTTACTACCTCAGTCAGTCACACAAATGAA 2774

Qy      396 TGGGATT-----ACAGGCGTGCAGCCGTGCCTGGCCTCTTTTAAATTTTATTTTTTTTAAGAC 450
            ||||| |     ||||| |||||  | || || ||
Db     2775 AAGGGTTTTTTCCCAATAAAATAGCAATTATTTTTGTTTGAGACAGCGTCTTTTTT 2834
```

RESULT 13
US-09-880-107-3428
Sequence 3428, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scheri, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02

```
Db     2775 AAGGGTTTTTTCCCAATAAAATAGCAATTATTTTTGTTTGAGACAGCGTCTTTTTT 2834
Qy      451 ACAGGGTGTCACACTCTTTACCCAGGATGAAGTTGCAGTGGTGTGGTGTTGTTTAAAAATTTTTTTT 510
Db     2835 GAGAGTCTCACTCTGTTTGCCCAGGCTGGAGTGCAGTGGCACGAATCTTGGCTCACTGCAAC 2894

Qy      511 CTTCAACTCCTGAGATCAAGC-ATCCTCCTCCTCAGCCTCCCAAGTAGCTGGGACCAAA 569
Db     2895 CTCTGCCTCCCGGGATCAAGCGAATTTTTCCTGCCTCACCCTCCCAAGTAGCTGGGA-TTAC 2953

Qy      570 GACATGCGACCACTACACCTGGCTAATTTTTTATTTTTAGTAGAGACGGGGT 629
Db     2954 GGTGTGTGCCACCACGCCCGGCTAATTTTTTGTATTTTTTAGTAGAGACGGGGT 2998

Qy      630 CTCAACTCTGTCACCCAGGCTGGTCTCGAACTCCTGACCTCG 689
Db     2999 CTT-GCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCAATCTCGG 3057

Qy      690 CCTCCCGGGTTCAAGTTTATTCTCCTGCCCCAGCCTCCCGAGTAGCTGGGACTACAGGCGC 749
Db     3058 CCTCCCAGGTTCATGCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCAT 3117

Qy      750 CCACCAGCCTAGCTAAT-TTTTTTGTATTTTTAGTAGAGATGGGGT----CACCATGT 804
Db     3118 GTATCACTACGCCCAGGTAATTTTTGTATTTTTAGTAGACAGGGGTTTCACCATGTTGG 3177

Qy      805 TCGCCAGGTTGATCTCGATCTCCTGGACCTTGTGATCCTGCCTGCCTCCCAAAGTG 864
Db     3178 GGGTCAGGCTGGTCTTGAACTCCTCGTGATCCTGCCTCGGCCTCCCAAAGTG 3237

Qy      865 CTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTATTTTTAATTTTTTTGTTTGTTTGAAA 924
Db     3238 CTGGGATTACCGGGCGTGAGCCACCGCGCCCGGCCT-----AATTTTTTTTTTGTTTTTGAGA 3290

Qy      925 TGGGAATCTCACTCTGTTTACCCAGGCTGGAGTGCAAATGGCCAAAATCTCGGCTCACTGCAAC 984
Db     3291 CAGGATCTCACTCTTCATCCAGGCTGAGAGTGCAGTGGTGCATCTCGGCTCACTGCAAC 3350

Qy      985 CTCTGCCTCCCGGGTTCAAGCGATTCTCCTGTGTCTCAAGGCAGAGTGCGAGTAGCTGGGATTACG 1044
Db     3351 CTCTGCCTCCCGGGTTCAAGCAAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACA 3410

Qy     1045 GGCACCTGCCACCACCACCCCGCTAATT-TTTGTATTTTCATTAGAGGGGGGGTTTCACCA 1103
Db     3411 GGCGCGCGCCACCATGCCCGGCTAATTTTTTGTATTTTTAGTAGAGATAGGGGTTTCACCA 3470

Qy     1104 TATTTGTCAGCGTGGTTCTCAAAATCCTGACCTCAGGTGATCCACCCGCCTCAGCCTCCCA 1163
Db     3471 CATTGGCCAGGCTGGTCTCAAAACTCTTGACCTCAAGTGATCACCACCCTCCTTGGCCTCCCA 3530

Qy     1164 AAGTGCTGGGATTACAGGCGTGAGCCACCTCACCCAGCCGGCT 1206
Db     3531 AAGTGCTGGGATTACAGGCGTGGGCCACTGCACCTGGCCTGTT 3573
```

```
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 3428
;   LENGTH: 99014
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
US-09-880-107-3428

Query Match   34.0%;  Score 490.4;  DB 10;  Length 99014;
Best Local Similarity  69.9%;  Pred. No. 1.2e-93;
Matches  849;  Conservative  0;  Mismatches  316;  Indels  49;  Gaps  12;

Qy      1 TTTTTTTTTTTTTTTTGAGATGGAGTTTTGTTTTGTTGTGAAATGGAATTCACTC-TGTTACCCAGGCTG 60
          ||||| |   | |||| ||||| |||||| |||||||                 |      |||||||||||
Db  88970 TTTTTCTTCTTCTTTTTTGAGACAGTTTCTGCTCTGTTGTCCCAGGCTGGAGTGCAATGGCGTGGCGA 89029

Qy     61 TCTCAGCTCACCGCACCATTCTCCCGGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89030 TCTTGGCTCACCGCACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 89089

Qy    121 CAGTAGCTGGGATTACAGGCATGCACCACCACGCCCGGCTAATTTTTGTATTTTTTTAG 180
          |||||||||||||||||||||||||||||||||||||||||     |||||||||||||
Db  89090 GAGTAGCTGGGATTACAGGCATGCACCACCACGCCCAGCTAATTTTTGTA----TTTTTGG 89145

Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAGGCTCAGGTGATC 240
          ||||||||| |||| |||| |||||||||||||||||||||||||||||||||
Db  89146 TAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 89205

Qy    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCCT-CTGC 299
          ||||| ||||||||||||||||||||||||||||| |||||||||||||||||||  ||||
Db  89206 GGCCACGCCCTGGCACTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATGCCCGGCC 89265

Qy    300 CTGGCTAATTTTTGTGTAGTAGAAACACAGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAG 359
          |||||||||||||||||||||   |     ||||  ||||||||||||||||||| |||||||
Db  89266 CTGGAATATCTTTTGTTGGTAGATATTATTGTTGGATT----------TGCTTAC 89303

Qy    360 CTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCACCGTG 419
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89304 TTGTATTTGTTAAGGACTTCAGCATCTACTTCATTAGGGGATAATTGGTCGTAGTTTG 89363

Qy    420 CCTGGCCTTTTTAATTTTTATTTTTATTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGATGAA 479
          ||||||||| | |||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89364 TTTTTGGTTATGTGCTTCTGCTTTGGTATTAGAATGATAGTGGCTTCATAGA--AT 89421

Qy    480 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACCTCCTGAGATCAAGCCATCCTCCT 539
          |||||||||||||||||||||||||||||||||| | |||||||||||||||||||||||||
Db  89422 GATTTAGGAGAGGTTCCCCTCCTTCCCCTTATCCTTGTGGAAATAGTGTCAATACAAATTCTTCTT 89481

Qy    540 GCCTCAGCCTCCCAAAGTGACTGGGACCAAAGACATGCACCACTTACACCTGGCTAATTTTT 599
          ||||||||||||||||||| ||||||||||||||||||||||     |||||||||||||||
Db  89482 TGAATGTCTGGTAGAAATCGCTGCGTGAATCCGTGGGACTTTGTGTTTTTGGTTG 89541

Qy    600 ATTTTTTATTTTTTAATTTTTTGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAG 659
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  89542 TTAATTTTTTTTATTTTTTTGAGACAGAGTCTT-GCTCTGTCGCCCAGGCTGGAGTGCAG 89600

Qy    660 TGGCGCAATCTTGGCTCACTGCAAGCTCC-GCTCTGCTCACCAGGCTGGAGTGCAG 719
          |||||||||||||||||||||||||||||  ||| |||||||||||||||||||||
Db  89601 TGGTGCAATCTCAGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCCTC 89660

Qy    720 AGCCTCCTGAGTAGCTGGGACTACAGGCGC---CCACCACGCCCGGCTAATTTTTGTA 776
          ||||||||| || ||||||||||||| |     |||||||||||||||||||||| 
Db  89661 AGCCTCCTGAGTTACAGGCGTGCCACCACACTCGGCTAATTTTTGTATTTTTTAG 89720

Qy    777 TTTTTTAGTAGAGATGGGG-TTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCG 835
          |||||||| |||| ||||| ||| ||||| |||||||||||||||||| ||||||||| ||
Db  89721 TTTTTAGTAGAGACGGGGTTTCACCATATTGGCCAGGCTGGTCTTGAACTCCTGACCTCG 89780

Qy    836 TGATCTGCCTGCCTCGGCCTCCCA---AAGTGCTGGGATTACAGGCGTGAGCCACCACGC 892
          ||||||||| ||||||||||||||   |||||||||||||||||||||||||||||||||
Db  89781 TGATCCACTCGGCCTCGCCTTGGTCTTCAAAGTATATGCTGGGATTACAGGCGTGAGCCACCACGC 89840
```

```
Qy    893 CCGGCCTTATTTTTAATTTTTGTTTTGTTGTGAAATGGAATTCACTC-TGTTACCCAGGCTG 951
          |||| |                                            ||||||||||||
Db  89841 CTGGC-----------CCTTTTTTGTTTTGTTTTGAGATGGAGTCTCACTCTGTTGCCCAGGCTG 89892

Qy    952 GAGTGCAATGGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCT 1011
          |||||||| |||||||||| ||||||||| ||||||||||||||||||||||||||||||
Db  89893 GAGTGCAGTGGCTCAATCTCAGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCT 89952

Qy   1012 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACAGGCGCCCGCCACCACACCCGGCTAA--1069
          ||||| ||||||||||||| |||||||||||||||| |||| |||||||||||||||||
Db  89953 CCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGTGCCCGCCACCACGCCCAGCTAATG 90012

Qy   1070 TTTTTGTATTTTTTTCATTTAGTAGAGACGGGGTTTCACCATATTT-GTCAGGCTGGTCTCAAACTC 1128
          ||||||| ||| ||||| |||| | ||||||||||||||||||||   ||||||||||||||||||||
Db  90013 TTTTTGTATTTTTTACTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTC 90072

Qy   1129 CTGACCTCAGGTGACCCACCTGCCTCGGCCTCCCAAGTGCTGGGATTACAGGCGTGAGC 1188
          |||||||||||||| ||||||| |||||||||||||||||||||||||||||||||||
Db  90073 CTGACCTCAGGTGATCCATCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGC 90132

Qy   1189 CACCTCACCCAGCC 1202
          |||| | ||||||
Db  90133 CACCACGCCCAGCC 90146
```

```
RESULT 14
US-09-918-686-2
Sequence 2, Application US/09918686
Patent No. US20020076720A1
; GENERAL INFORMATION:
;   APPLICANT: Brunkow, Mary
;   APPLICANT: Proll, Sean
;   APPLICANT: Paeper, Bryan
;   APPLICANT: Staehling-Hampton, Karen
;   TITLE OF INVENTION: METHODS FOR IDENTIFYING
;   TITLE OF INVENTION: GENOMIC DELETIONS
;   FILE REFERENCE: 240083.515
;   CURRENT APPLICATION NUMBER: US/09/918,686
;   CURRENT FILING DATE: 2001-07-30
;   NUMBER OF SEQ ID NOS: 105
;   SOFTWARE: FastSEQ for Windows Version 4.0
US-09-918-686-2

; SEQ ID NO 2
;   LENGTH: 51719
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1246, 2572, 2604
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match   33.2%;  Score 479.2;  DB 10;  Length 51719;
Best Local Similarity  68.3%;  Pred. No. 2.3e-91;
Matches  850;  Conservative  0;  Mismatches  358;  Indels  36;  Gaps  12;

Qy      1 TTTTTTTTTTTTTTGGAGATGGAGTTTTGTTTGTTGTGAAATGGAATTCACTC-TGTTACCCAGGCTG 60
          |||||||||||||| |||||||||||||| |||||| |||                 |||||||||||
Db  18838 TTTTTTTTTTTTTTTGAGACGGAGTTTT-GCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCGA 18896

Qy     61 TCTCAGCTCACCGCAACCTCCGGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          ||||||||||||||| ||||   ||||||||||||||||||||||||||||||||||||||
Db  18897 TCTCGGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 18956

Qy    121 CAGTAGCTGGGATTACAGGCATGCACGTCACCACGCCCGGCTAATTTTTGTATTTTTTTAG 180
          ||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||| ||
Db  18957 GAGGAGCTGGGATTACAGGCATGTGCCACCACTACCAGGCCGGCTAATTTTTGTATTTTTTAG 19012

Qy    181 TAGAGATGGGGTTTCTTCGTGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCGTGATC 239
          |||||| ||||||||| || |||||||||||||||||||||||||||||||||| |||
Db  19013 TAGAGATGGGGTTTCTTCGTTGGTCTTCAAGCTGGTGGTCTTGACCTCAGGTGATC 19072

Qy    240 ----CCCTCGTCTCGGCCTCCCAAAGTGCTAG--ATACAGGACTGGCCACCATGCCCG 293
              |||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db          ... 293
```

```
; US-09-918-686-1
; Sequence 1, Application US/09918686
; Patent No. US20020076720A1
; GENERAL INFORMATION:
;  APPLICANT: Brunkow, Mary
;  APPLICANT: Proll, Sean
;  APPLICANT: Paeper, Bryan
;  APPLICANT: Staehling-Hampton, Karen
;  TITLE OF INVENTION: METHODS FOR IDENTIFYING
;  TITLE OF INVENTION: GENOMIC DELETIONS
;  FILE REFERENCE: 240083.515
;  CURRENT APPLICATION NUMBER: US/09,918,686
;  CURRENT FILING DATE: 2001-07-30
;  NUMBER OF SEQ ID NOS: 105
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  OTHER INFORMATION: n = A,T,C or G
; US-09-918-686-1
;
;  SEQ ID NO 1
;  LENGTH: 92139
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: 7043, 8369, 8401
;  OTHER INFORMATION: LOCATION: n = A,T,C or G

RESULT 15

Query Match      33.2%;  Score 479.2;  DB 10;  Length 92139;
Best Local Similarity   68.3%;  Pred. No. 2.6e-91;
Matches  850;  Conservative  0;  Mismatches  358;  Indels  36;  Gaps  12;
```

(Sequence alignment data — Qy/Db aligned DNA sequences, positions as printed)

```
Qy    706  TATTCTCCTGCCCACCCTCCTGAGTAGTAGCTGGGACTACAGGCG---CCCACCACGCCTAG  762
           || ||| |||||||||||||||||||||||||||||||| ||||| ||||| || ||||||
Db  25338  AATCCTCTCACCTCAGCCCTTTGAGTAGCTGGGACCACAGGTGTGTTACCACACCCAG  25397

Qy    763  CTAATTTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATCTTGA  822
           ||||||||| | |||||||||||||||||||||||||||||||| |||||||||||||||
Db  25398  ATAATTTTTTGTGCTTTTTGTAGAGACAGGGTTTCGACAGTTGCCCAGGCCGGACTAAAA  25457

Qy    823  TCTCTCGGAC-CTTGTGATCTGCCTCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG  881
           | |||||| ||||||||||||||||||||||||||||||||||||||||||||||| ||
Db  25458  TTCCTCGGGCTCAAGTGATCCTCCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGTATG  25517

Qy    882  AGCCACCACGCCCGGCTTATTTTTAATTTTTTGTTTGTTGAAATGGAATCTCACTCTGTT  941
           ||||| || | ||||||||||||||||||||||||||||||||||||||||||| ||||
Db  25518  TGCCATCGTCACCAGCCAGAATATGCATTTTCTTTTTTTCAGACGGGGTCTCACTCTGTT  25577

Qy    942  ACCCAGGCTGGAGTGCAAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTC  1001
           ||||||||||||||||||||| |||||||||||||||||||||||||||| | ||||||
Db  25578  GCCCAGGCTGGAGTGCAGTGGTGATCTCAGCTCACTGCAAACCTCTGCCTTCCAGGTTC  25637

Qy   1002  AAGCGATTCTCCTGTCTCAGCCTCCCAAGCACCTGGGATTACAGGCGCACCTGCCACCACAC  1061
           ||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
Db  25638  AAGCGATTCTCCTGCCTCAGCCTCCCAAGCACCTAGGATTACAGGCGTGCTGCCACACAC  25697

Qy   1062  CCCGGCTAATTTTTGTATTTTCATTAGAGACGGGGGTTTCACCATATTGTTCAGGCTGGTCT  1121
           |||||||||||||| ||||||||||||||||| ||| |||||| | ||||||||||||||
Db  25698  CCGGCTAATTTTTGCATTTTTAGTAAAGATGGGGTTTCACCATGTTGGTCAGGCTGATCA  25757

Qy   1122  CAAACTCCTGACCTCAGGTGACCCACCCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGG  1181
           |||||||| ||| ||||| |||||||||||||||||||||||||||||||||||||||||
Db  25758  CAAACTCCCGACTGCAAGTGATCCGCTTACCTCGGCCTCCCAAAGTGTTAGGATTACAGG  25817

Qy   1182  CGTGAGCCACCTCACCCAGCCGCCTAATTTAGATAAAAAAATAT  1225
           | ||||||||||| ||||| | |||| || |||
Db  25818  CATGAGCCATTGTGCCTGGCCGAATATGCCACTTTTAATAAGCAT  25861
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 1, 2003, 21:20:53 ; Search time 66 Seconds
                   (without alignments)
                   6700.423 Million cell updates/sec

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 ttttttttttggagtggag...........ttaaacaaagcttagagca 1442

Scoring table: IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                   1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                   2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                   3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                   4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                   5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                   6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	% Match	Length DB	DB ID	Description
1	1442	100.0	2	1442	US-08-454-557C-120	Sequence 120, App
2	1442	100.0	2	1442	US-08-340-426D-120	Sequence 120, App
3	1442	100.0	2	1442	US-08-450-673C-120	Sequence 120, App
4	1223.4	100.0	5	1418	PCT-US95-17111A-120	Sequence 120, App
5	1080.2	74.9	2	1381	US-08-454-557C-49	Sequence 49, Appl
6	1080.2	74.9	2	1381	US-08-340-426D-49	Sequence 49, Appl
7	1080.2	74.9	2	1381	US-08-450-673C-49	Sequence 49, Appl
8	1080.2	74.9	5	1381	PCT-US95-17111A-49	Sequence 49, Appl
9	475.4	33.0	2	14796	US-08-975-080-35	Sequence 35, Appl
10	475.4	33.0	2	14796	US-09-630-706-10	Sequence 10, Appl
11	475.4	33.0	4	14796	US-09-496-694B-3	Sequence 3, Appl
12	416.4	28.9	2	59065	US-09-813-817-3	Sequence 3, Appl
13	416.4	28.9	5	59065	US-09-978-197-3	Sequence 3, Appl
14	412.2	28.6	4	43950	US-09-735-934A-3	Sequence 3, Appl
15	400.8	27.8	2	5543	US-08-687-080-101	Sequence 101, Appl
16	393.6	27.3	2	4421	US-08-367-963B-9	Sequence 9, Appl
17	393.6	27.3	4	4421	US-08-257-841A-9	Sequence 9, Appl
18	393.6	27.3	5	4421	PCT-US95-0201-9	Sequence 9, Appl
19	393	27.3	2	4421	US-08-658-136-1	Sequence 1, Appl
20	393	27.3	3	53526	US-08-658-136-2	Sequence 2, Appl
21	393	27.3	3	53577	PCT-US99-0201-9	Sequence 9, Appli
22	392	27.2	4	62804	US-09-800-960-3	Sequence 3, Appli
23	386.4	26.8	4	43950	US-09-735-934A-1	Sequence 1, Appli
24	386	26.8	1	31571	US-08-323-443B-1	Sequence 10, Appl
25	376.8	26.1	2	99500	US-09-798-096-10	Sequence 10, Appl
26	374.6	26.0	4	11636	US-09-173-914-6	Sequence 6, Appl
27	374.4	26.0	4	2598	US-09-026-033-18	Sequence 18, Appl
28	371.4	25.8	4	26664	US-09-564-805-28	Sequence 28, Appl
29	370	25.7	4	4793	US-09-561-497-10	Sequence 10, Appl
30	367.4	25.5	1	6769	US-08-480-784-20	Sequence 20, Appl
31	367.4	25.5	1	6769	US-08-483-553-20	Sequence 20, Appl
32	367.4	25.5	1	6769	US-08-487-002-20	Sequence 20, App
33	367.4	25.5	1	6769	US-08-483-554B-20	Sequence 20, Appl
34	367.4	25.5	1	6769	US-08-488-011B-20	Sequence 20, Appl
35	367.4	25.5	1	6769	US-08-850-727-20	Sequence 20, Appl
36	367.4	25.5	5	6769	PCT-US95-10202-20	Sequence 20, Appl
37	367.4	25.5	5	6769	PCT-US95-10203-20	Sequence 20, Appl
38	367.2	25.5	1	8392	US-08-080-255-6	Sequence 6, Appl
39	357.2	24.8	1	8392	US-08-465-713-6	Sequence 6, Appl
40	357.2	24.8	5	8392	PCT-US93-05857-6	Sequence 6, Appl
41	356.8	24.7	4	84495	US-09-797-906-3	Sequence 3, Appl
42	356.8	24.7	1	84495	US-08-488-011B-20	Sequence 20, Appl
43	356.6	24.7	4	45716	US-09-965-048-5	Sequence 5, Appli
44	356.6	24.7	4	45989	US-08-965-048-6	Sequence 6, Appli
45	355.8	24.7	1	3035	US-08-726-725-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-454-557C-120
Sequence 120, Application US/08454557C
Patent No. 5830670

GENERAL INFORMATION:
  APPLICANT: de la Monte, Suzanne
  APPLICANT: Wands, Jack R.
  TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
  TITLE OF INVENTION: of Alzheimer's Disease
  NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:
  ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
  STREET: 1100 New York Avenue, Suite 600
  CITY: Washington
  STATE: D.C.
  COUNTRY: U.S.A.
  ZIP: 20005-3934

COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/454,557
  FILING DATE: 30-MAY-1995
  CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
  NAME: Ludwig, Steven R.
  REGISTRATION NUMBER: 36,203
  REFERENCE/DOCKET NUMBER: 0609.3840003

TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 371-2600
  TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
  LENGTH: 1442 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:
  NAME/KEY: CDS
  LOCATION: 15..1139

US-08-454-557C-120

Query Match   100.0%; Score 1442; DB 2; Length 1442;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TTTTTTTTTTTTGAGATGGAGTTTTTCGCTCTTGTTTCCCAGGCTGGAGTGCAATGGCGAA 60

```
Db     1 TTTTTTTTTTTGAGATGGAGTTTTCGTCTTGTTGCCCAGGCTGGAGTGCGCAA       60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     1 ...........................................................      60

Qy    61 TCTCAGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC     120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTCAGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC     120

Qy   121 CAGTAGCTGGGATTACAGGCGCATGCCACCACGCCCGGCTAATTTTTGTATTTTTTAG      180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTAGCTGGGATTACAGGCGCATGCCACCACGCCCGGCTAATTTTTGTATTTTTTAG      180

Qy   181 TAGAGATGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGATC     240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGTGATC     240

Qy   241 CCTCCGCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGGTCACCGCGCCCGGCC    300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCGCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGGTCACCGCGCCCGGCC    300

Qy   301 TGGCTAATTTTTGTAGAGACAGGGTTTCACCGTGTTAGCCAAGCTGGTCTCGAACTCCTGAGC   360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGCTAATTTTTGTGGTAGAGACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCGAACTCCTGAGC   360

Qy   361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC    420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC    420

Qy   421 CTGGCCTTTTTTATTTTATTTTTTTTTTTTAAGACACAGGGTGTCCCACTCTCTTACCCAGGATGAAG   480
         |||||||||||||||||||||||||||||||||||||||||
Db   421 CTGGCCTTTTTTATTTTATTTTTTTTTTTAAGACACAGGGTGTCCCACTCTCTTACCCAGGATGAAG   480

Qy   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAATCACCCTCCTG    540
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAATCACCCTCCTG    540

Qy   541 CCTCAGCCTCCCAAAGTAGCTGGGACCAAAAGACATGCACCACCACTACCAGCTGGCTAATTTTTA    600
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAGCCTCCCAAAGTAGCTGGGACCACAAAGACATGCACCACCACTACCAGCTGGCTAATTTTTA    600

Qy   601 TTTTTTATTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT     660
         |||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTTTTATTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT     660

Qy   661 GGCGCAAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA     720
         ||||||||||||||||||||||||||||||||||||||||||||
Db   661 GGCGCAAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA     720

Qy   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCGCCACCACGCCCTAGCTAATTTTTTGTATTTTT     780
         |||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCGCCACCACGCCCTAGCTAATTTTTTGTATTTTT     780

Qy   781 TAGTAGAGATGGGGTTCACCATGTTCGCCAGGCTGATCTTGAATCTCTGGACCTTGTGATC     840
         |||||||||||||||||||||||||||||||||||||||||||
Db   781 TAGTAGAGATGGGGTTCACCATGTTCGCCAGGCTGATCTTGAATCTCTGGACCTTGTGATC     840

Qy   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGGCTTA     900
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGGCTTA     900

Qy   901 TTTTTTAATTTTGTTTGTTTGTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT     960
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTTTTAATTTTTGTTTGTTTGTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT     960

Qy   961 GGCCAAAATCTCGGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGTCTCA    1020
         ||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCCAAAATCTCGGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTGTCTCA    1020

Qy  1021 GCCTCCCAAGCAGCTGGGATTACGGGCACCACCGCACCACCCGCCTAATTTTTGTATTTT    1080
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GCCTCCCAAGCAGCTGGGATTACGGGCACCACCGCACCACCCGCCTAATTTTTGTATTTT    1080

Qy  1081 TCATTAGAGGCGGGGTTTCACCATATATTTGTCAGGCTGGTCTCAAAACTCCTGACCTCAGGT    1140
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TCATTAGAGGCGGGGGTTTCACCATATATTTGTCAGGCTGGTCTCAAAACTCCTGACCTCAGGT    1140

Qy  1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCACCTCACCCAG    1200
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCACCTCACCCAG    1200

Qy  1201 CCGGCTAATTTGATAAAAAAAAATATGTAGCAATGATGTGGGGGGGGGTCTTGCTATGTTGCCCAGGCT    1260
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CCGGCTAATTTAGATAAAAAAAAATATGTAGCAATGATGGGGGGGTCTTGCTATGTTGCCCAGGCT    1260

Qy  1261 GGTCTCAAAACTTCTGGCTTCATTGCAATCCTTCCAAATGAGCCACCACACAGCCAGCTCA    1320
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GGTCTCAAAACTTCTGGCTTCATTGCAATCCTTCCAAATGAGCCACCACACAGCCAGCTCA    1320

Qy  1321 CATTTTTTTAAACAGTTTACATCTTTATTTTAGTATACTTAGTAATACAATAAACATTGT    1380
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CATTTTTTTAAACAGTTTACATCTTTATTTTAGTATACTTAGTAATACAATAAACATTGT    1380

Qy  1381 CAAACCTGCAAATTCAGTAAACAGAGTTCTTTTATTAACTTTAAACAAAGCTTTAGAG    1440
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAACCTGCAAATTCAGTAGTAAACAGAGTTCTTTTATTAACTTTAAACAAAGCTTTAGAG    1440

Qy  1441 CA    1442
         ||
Db  1441 CA    1442
```

Query Match   100.0%;   Score 1442;   DB 2;   Length 1442;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1442; Conservative 0;

```
Qy      1 TTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTTGCCCAGGCTGGAGTGCAATGGCGCGAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTTGCCCAGGCTGGAGTGCAATGGCGCGAA 60

Qy     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy    121 CAGTAGCTGGGATTACAGGCACGCACCACCGCGCCCGGCTAATTTTTGTATTTTTAGTAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCACGCACCACCGCGCCCGGCTAATTTTTGTATTTTTAGTAG 180

Qy    181 TAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240

Qy    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGCGTCCCGACCTCAGATGATCAGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGCGTCCCGACCTCAGATGATCAGCC 300

Qy    301 TGGCCTAATTTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCCTAATTTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC 360

Qy    361 TCAAGCAGTCCACCTGTCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAGTCCACCTGTCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTGC 420

Qy    421 CTGGCCTTTTATTTTATTTTTTTAGACACAGAGTGTCCCACTCTTACCCAGGATGAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGCCTTTTATTTTATTTTTTTAGACACAGAGTGTCCCACTCTTACCCAGGATGAAG 480

Qy    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCTCCTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCTCCTG 540

Qy    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACACCTGGCTAATTTTTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACACCTGGCTAATTTTTA 600

Qy    601 TTTTTTAATTTTTTAATTTTTTGAGACAGAGTCTCAACTCGTCACCCAGGCTGGAGTGCAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTTAATTTTTTAATTTTTTGAGACAGAGTCTCAACTCGTCACCCAGGCTGGAGTGCAGT 660

Qy    661 GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCCA 720

Qy    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCCCACCTAGCTAATTTTTGTATTTTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCCCACCTAGCTAATTTTTGTATTTTT 780

Qy    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCCTGACCTTGTGATC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCCTGACCTTGTGATC 840

Qy    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGGCTTA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGGCTTA 900

Qy    901 TTTTTTAATTTTGTTTGTTTGTTTGAAATGGAGTCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTTTAATTTTGTTTGTTTGTTTGAAATGGAGTCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960

Qy    961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020

Qy   1021 GCCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACCACCACCCGGCTAATTTTTGTATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACCACCACCCGGCTAATTTTTGTATTT 1080

Qy   1081 TCATTAGAGGCGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140

Qy   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACCTCACCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACCTCACCCAG 1200

Qy   1201 CCGGGCTAATTTAGATAAAAAAAATATGTAGCAAATGGGGGTCTTGCTATGTTGCCCAGGCT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGGCTAATTTAGATAAAAAAAATATGTAGCAAATGGGGGTCTTGCTATGTTGCCCAGGCT 1260

Qy   1261 GGTCTCAAAACTTCTGGCCTTCATGCAAATCCTTCCAAATGAGGCCACAACACCAGCCAGTCA 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCAAAACTTCTGGCCTTCATGCAAATCCTTCCAAATGAGGCCACAACACCAGCCAGTCA 1320

Qy   1321 CATTTTTTAAACAGTTACATCTTTATTTTAGTATATACTAGAAAGTAATACAATAGAACATGT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTTAAACAGTTACATCTTTATTTTAGTATATACTAGAAAGTAATACAATAGAACATGT 1380

Qy   1381 CAAACCTGCAAAATCAGTAGTAACAGAGTTCTTTTTATAACTTTTTAAACAAAAGCTTTTAGAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAAATCAGTAGTAACAGAGTTCTTTTTATAACTTTTTAAACAAAAGCTTTTAGAG 1440

Qy   1441 CA 1442
          ||
Db   1441 CA 1442
```

RESULT 3
US-08-450-6673C-120
Sequence 120, Application US/08450673C
Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne
APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 120:

SEQUENCE CHARACTERISTICS:
LENGTH: 1442 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS

```
;   LOCATION: 15..1139
US-08-450-673C-120

Query Match    100.0%;      Score 1442;   DB 2;   Length 1442;
Best Local Similarity 100.0%;      Pred. No. 0;
Matches 1442;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1  TTTTTTTTTTTTTGAGATGGAGAGTTTTCGCTCTGTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA      60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTTTTTTTTTTGAGATGGAGAGTTTTCGCTCTGTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA      60

Qy     61  TCTCAGCTCACCGCAACCTCCCGCCTCCCGGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC     120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  TCTCAGCTCACCGCAACCTCCCGCCTCCCGGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC     120

Qy    121  CAGTAGCTGGGATTACAGGCATGCGCCACCACGCCTCGGCTAATTTTTGTATTTTTTAG     180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CAGTAGCTGGGATTACAGGCATGCGCCACCACGCCTCGGCTAATTTTTGTATTTTTTAG     180

Qy    181  TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATC     240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATC     240

Qy    241  CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCGCGCCCGGCCTGCC     300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCGCGCCCGGCCTGCC     300

Qy    301  TGGCTAATTTTTGTAGTAGAAAACAGGGTTTCACTGATGTGCCCAGGCTGGTCTCCTGAGC     360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  TGGCTAATTTTTGTGGTAGAAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC     360

Qy    361  TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC     420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC     420

Qy    421  CTGGCCTTTTTATTTTTATTTTTTGAAGACAGGGTCCCACTCTTACCCAGGATGAAG     480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTGGCCTTTTTATTTTTATTTTTTGAAGACAGGGTCCCACTCTTACCCAGGATGAAG     480

Qy    481  TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCCTCCTG     540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCCTCCTG     540

Qy    541  CCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTA     600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  CCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTA     600

Qy    601  TTTTTATTTTTGTAGAGACAGTCACCTCTCGTCACCCAGGCTGGAGTGCAGT     660
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TTTTTAATTTTTGTAGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT     660

Qy    661  GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA     720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA     720

Qy    721  GCCTCCTGACTAGCTGGGACTACAGGCGCGCCACCACGCCCGGCTAATTTTTGTATTT     780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GCCTCCTGACTAGCTGGGACTACAGGCGCCCAGCTAATTTTTTGTATTTTTTAGTATTT     780

Qy    781  TAGTAGAGATGGGGTTTCACCATGTTGCCCAGGCTGGTCTCGATCTCCTGGACCTCGTGATC     840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  TAGTAGAGATGGGGTTTCACCATGTTGATCTTGGACCTCGTGATC     840

Qy    841  TGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCACGGCCGGCCTTA     900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCACGGCCGGCCTTA     900

Qy    901  TTTTTAATTTTTTGTTTTGTTTGGAGATCACTTCACTGCTGTTACCCAGGCTGGAGTGCAAT     960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TTTTTAATTTTTTGTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT     960

Qy    961  GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA     1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA     1020

Qy   1021  GCCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACCACACCCGGCTAATTTTTGTATTT     1080
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  GCCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACCACACCCGGCTAATTTTTGTATTT     1080

Qy   1081  TCATTAGAGGCGGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT     1140
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TCATTAGAGGCGGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT     1140

Qy   1141  GACCCACCTGCCTCAGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCACCCAG     1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCACCCAG     1200

Qy   1201  CCGGCTAATTTTAGTAAAAAAATATGTAGCAAATGGGGGGTTGTTGCTATGTTGCCCAGGCT     1260
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CCGGCTAATTTTAGTAAAAAAATATGTAGCAAATGGGGGGTTGTTGCTATGTTGCCCAGGCT     1260

Qy   1261  GGTCTCAAACTTCTGCCTTCATGCAATCCTTCCAAAATGAGCCACAACACCCAGCCAGTCA     1320
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  GGTCTCAAACTTCTGGCCTTCATGCAATCCTTCCAAAATGAGCCACAACACCCAGCCAGTCA     1320

Qy   1321  CATTTTTTAAACAGTTACATCTTTATTGTAGTAGTAACTAGAAGTAATACAATAAACATGT     1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  CATTTTTTAAACAGTTACATCTTTATTGTAGTAGTAACTAGAAGTAATACAATAAACATGT     1380

Qy   1381  CAAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTAAACTTTTTAAAACAGTTTTGAG     1440
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  CAAAACCTGCAAAATTCAGTAGTAACAGAGTTCTTTTAAACTTTTTAAAACAGTTTTGAG     1440

Qy   1441  CA     1442
           ||
Db   1441  CA     1442
```

```
TYPE:         nucleic acid
STRANDEDNESS: both
TOPOLOGY:     both
FEATURE:
  NAME/KEY:   CDS
  LOCATION:   14..1418

PCT-US95-17111A-120

Query Match    84.8%;  Score 1223.4;  DB 5;  Length 1418;
Best Local Similarity  96.9%;  Pred. No. 0;
Matches 1375;  Conservative  0;  Mismatches  31;  Indels  13;  Gaps  12;
```

```
Qy      2 TTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTCTTGTTCCCAGGCTGGAGTGCAAT    61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTTCCCAGGCTGGAGTGCAATGGCGCAAT    60

Qy     62 CTCAGCTCACCGGACCTCCGCGTTCAAGCGATTCTCCTGCCTCAGCCTCCC    121
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC    120

Qy    122 AGTA-GCTGGGATTACAGGCATGCGACCACCACGCCTGGGCTAATTTTTGTATTTTTTTTAG    180
          ||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTAGGCTGGGATTACAGGCATGCGA--CCACGCTCGGCTAATTTTTGTATTTTTTTTAG    179

Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGGAACTCCGACCTCAGATGATC    240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACT-CCGACCTCAGATGATC    238

Qy    241 CCTCCGCTCGGCCTCCCAAAGTGCTAGATTACAGGACTGAGCCACCATGCCCGGCCTGC    299
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    239 CTCCCGCTCGGCCTCCCAAAGTGCTAGATTACAGGCGTGAGCCACCATGCCCGGCCTGC    298

Qy    300 CTGGGCTAATTTTTGTTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAG    359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 CTGGCTAATTTTTGTTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAG    358

Qy    360 CTCAAGCGAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTG    419
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 CTCAAGCGAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTG    418

Qy    420 CCTGGGCCTTTTTAATTTATTTTTTATTTTTTTTAAGACACAGAGTTCCCACTCTTACCCAGGATGAA    479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 CCTGGGCCTTTTTATTTTTTATTTTTTTTAAGACACAGAGTTCCCACTCTTACCCAGGATGAA    478

Qy    480 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCAACTCCTGAGATCAAGCACATCCTCTCT    539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479 GTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCAACT-CTTGAGATCAAGCACATCCTCT    537

Qy    540 GCCTCAGCCTCCC-AAGTAGCTGGGACCAAAGACACTACACCACCATCTGGCTAATTTT    598
          ||||||||||||  |||||||||||||||||||||||||||||||||||||||||||||
Db    538 GCCTCAGCCTCCCAAAGTAGCTGGGACCACAAAGACACTACACCACCATCTGGCTAATTTT    597

Qy    599 TATTTTTATTTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCA    658
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TATTTTTATTTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCA    657

Qy    659 GTGGGGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTATTCCTCCTCGCC    718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 GTGGCGCAAATCTTGGCTCACTGCAACCTCCGGGTTCAAGTGATTCTCCTGCCC    717

Qy    719 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGGCCTAGCTAATTTTTGTATT    778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGGCCTAGCTAATTTTTGTATT    777

Qy    779 TTTAGTAGAGATGGGG--TTCACCATGTTGGCCAGGTTGAT-CTTGATCTCTGGACCTTGT    836
          |||||||||||||||  |||||||||||||||||||  |||||  |||||||||||||
Db    778 TTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATGCTAGATCTCTTGACCTTGT    837

Qy    837 GATCTGCCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG-CGTGAGCCACCAGCCCG    895
          ||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||
Db    838 GATCTGCCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGACGTGAGCCACCAGCGCCG    897

Qy    896 GCTTATTTTTTTAATTTGTTTGTTTGTTGAAATGAATGGAATCACTCTGTTTACCCAGGCTGGAGT    955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    898 GCCTATTTTTTAATTTGTTTGTTTGTTGAAATGGAATTCACTCTGTTTACCCAGGCTGGAGT    957

Qy    956 GCAATGGCCAAATCTGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTG    1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    958 GCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTG    1017

Qy   1016 TCTCAGCCTCCCCAAGCAGCTGGGATTACGGGCCACACACACCCCGGCTAATTTTTG    1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 TCTCAGCCTCCCCAAGCAGCTGGGATTACGGGCCACACACACCCCGGCTAATTTTTG    1076

Qy   1076 TATTTTCATTAGAGGCGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCT    1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 TATTTTCATTAGAGGCGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCT    1136

Qy   1136 CAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCA    1195
          |||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Db   1137 CAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGA--CGCCTCA    1194

Qy   1196 CCCAGCCCGGGCTAATTTAGTATAAAAAATAGTGAGCAATAGGGGGGGTCTTTGTATTGTTTGCCC    1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 CCCAGCCCGGGCTAATTTAGTATAAAAAATAGTGAGCAATAGGGGGGGTCTTTGTATTGTTTGCCC    1254

Qy   1256 AGGCTGGTCTCAAAACTCCTGGCCTTCATCGCAATCCTTCCAAATGAGCCACCACACACCAGCC    1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 AGGCTGGTCTCAAAACTCCTGGCCTTCATCGCAATCCTTCCAAATGAGCCACCACACACCAGCC    1314

Qy   1316 AGTCACATTTTTTAAACAGTTACACTCTTTAATTTTATTAGTATATACTAGAAGTAATAACAATAA    1375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 AGTCACA-TTTTAAACAGTTACACTCTTTAATTTTAGTATATACTAGAAAGTGATACGATAAC    1373

Qy   1376 CATGTCAAAACCTGCAAAATTCAGTAGTAGTACACAGAGTCTTTT    1414
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1374 ATGGCGGAACCTGCAAATTCGAGTAGTACACAGAGTCTTTT    1412
```

RESULT 5
US-08-454-557C-49
Sequence 49, Application US/08454557C
Patent No. 5830670
GENERAL INFORMATION:
  APPLICANT: de la Monte, Suzanne
  APPLICANT: Wands, Jack R.
  TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
  TITLE OF INVENTION: of Alzheimer's Disease
  NUMBER OF SEQUENCES: 121
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
  STREET: 1100 New York Avenue, Suite 600
  CITY: Washington
  STATE: D.C.
  COUNTRY: U.S.A.
  ZIP: 20005-3934
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/454,557C
  FILING DATE: 30-MAY-1995
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
  NAME: Ludwig, Steven R.
  REGISTRATION NUMBER: 36,203
  REFERENCE/DOCKET NUMBER: 0609.3840003
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 371-2600
  TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 49:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 1381 base pairs
  TYPE: nucleic acid

STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-49

Query Match 74.9%; Score 1080.2; DB 2; Length 1381;
Best Local Similarity 94.4%; Pred. No. 3.5e-292;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

```
Qy      2 TTTTTTTTTTGAGAATGGAGAGTTTTCGCTCTTGTTTCCCAGCGTGGAGTGGCAATGGCAAT 61
          |||||||||||||||||||| ||||||||||||||||| ||||||||||||||||||||||
Db      1 TTTTTTTTTTGAGAATGGAGAGTTTTCGCTCTTGTTTGCCCAGCGTGGAGTGGCGAAT      61

Qy     62 CTCAGCTCACCGCAACCTCCGGCGCTCCCGGGGTTCAAGGCGATTCTCCTCCTCAGCCTCCC 121
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGCAACCTCCGGCGCTCCCGGGGTTCAAGGCGATTCTCCTCCTCAGCCTCCC 120

Qy    122 AGTAGCTGGGATTACAGGCATGCACCACCACGCTCGGCTAATTTTTGTATTTTTTTTTAGT 181
          ||||||||||||||||||||||||||  ||||||||| |||||||||||||||||||||||
Db    121 AGTAGCTGGGATTACAGGCATGCA-CCACCTCGGCCTAATTTTTGTATTTTTTTTTAGT   179

Qy    182 AGAGATGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAAACTCCCGACCTCAGATGAT 239
          |||||||||||||  ||||||| |||||||||||||||||||||||||||||||||||| |
Db    180 AGAGATGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT 239

Qy    240 CCCTCCGTCTCGGCGCTCCCAAAGTGCT--AGATACAGGACTGGCCACCATGCCCGG-CT 295
          |||| ||||||||| ||||||||||||   ||||||||| |||||||||||||||||  ||
Db    240 CTCCCGTCTCGGCGTTCCCAAAGTGCTGAGATTACAGGCCATGAGCCACCATGCCCAGATT 239

Qy    296 CTGCCCTGGGCTAATTTTTGTGGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGG 354
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||| ||
Db    296 CTGCCCTGGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTG-TA 354

Qy    355 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAG 414
          |||||||||||||||| ||||||||| ||||||||||||||||||||||||||||| ||
Db    360 CTGAGCTCAAGCAGTCACCTGCCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG 418

Qy    415 CCGGTGCCTGGCCCTTTTTTTAATTTTTTTAAGACAGAGGGTCCCACTCTTACCCAGG 474
          ||||||||||| ||||||||||||||||||||||||||| |||||||||||||||||
Db    419 CCGTGCCTGGCCCTTTTTTATTTTTTAATTTTTGAGACAGTCTC-ACTCTGTCACCAGGCTGGA 478

Qy    475 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCGAGATCAAGC-AT 533
          |||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db    479 ATGAAGTGCAGTGGTGCGATCACAGCTCACTGCAGCCTTCAAACTCCTGGGGATCAAGCGAAT 538

Qy    534 CCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACCTCGCCTA 593
          |||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    539 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACCGCCTACACCTGG 597

Qy    594 ATTTTTTATTTTTATTTTTTGAGACAGAGTCTCACCCAGGCTGGA 653
          |||||||||||||||||||||||||||||||||||||||||||||
Db    598 ATTTTTTATTTTTATTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA 656

Qy    654 GTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCC 713
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    657 GTGCAGTGGGCCAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCC 716

Qy    714 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCAGCTAATTTTT 773
          |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    717 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCTAATTTTTT 773

Qy    774 GTATTTTTTAGTAGAGATGGGG-TTCACCATGTTGGTCAGGCTGGTCTCGAAC 832
          |||||||||||||||||||||  |||||||||||||||||||||||||||||
Db    777 GTATTTTTAGTAGAGACGGGGTTTCACCATGTTCGCCAGGTTGATCTCGAACTCC 836

Qy    833 TTGTGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGG 891
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    837 TTGTGACCTCGCCTGCCTCGGCCTCCCAAGTGCTGGGATTACAG--GTCGTGACTCCAC 894

Qy    892 CCCGGCGTTATTTTTTAATTTTTTTGTTTGTTTGAAATGGAAATCACTCTGTTACCCAGGCTG 951
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    895 GCCGGCCTATTTTTTAATTTTTTGTTTGTTTTGAAATGGAAATGGAATCTCGCCATGTTACCAGGCTG 954

Qy    952 GAGTGCAAATGGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGAATTCT 1011
          ||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    955 GAGTGCAATGG-CAAATCTCGGCTACTCGCAACCTCTGCCTCCCGGC-TCAAGCGGATTCT 1012

Qy   1012 CCTGTCTCAGCCTCCCAACACGACGTGGGAGTTACGGCGACCACCCCGCTAATT 1071
          ||||||||||||||||||  ||||| ||||||||||||| |||||| ||||| ||
Db   1013 CCTGTCTCAGCCTCCCAAGACAGCAGTCTGGGAGTTACGGC--ACCTGCACCACCACGCCCGGCTAATT 1071

Qy   1072 TTTGTATTTTCATTAGAGACGGGGGTTTCACCATATTTGTGCAGGCT-GGTCTCGAAAACTCCT 1130
          ||||||||||||||||||||||||||||||||||||| |||||  ||| |||||||  |||||
Db   1071 TTTGTATTTTCATTAGAGACGGGG--TTTACCATATTTGTCAGGCTGGTCTCGGGTGCTCTCAAAACTCCT 1130

Qy   1131 GACCTCAGGTGATCCACCACGCCTCCCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1190
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||| |||
Db   1131 GACCTCAGGTGACCTGCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1190

Qy   1191 CCTCACCCAGCGCGGGCTAATTTTGATAAAAAAAATATGTAGCAAATGGGGGTCTTGCTATGT 1250
          ||||||| |||| |||||||| |||| ||||||||| |||||||||||||  ||||||||||||
Db   1189 CCTCACCCAGCCCGGGCTAATTTTTGTGATANAAATATGTAGCAAATGTGTGGGGG--TCTGTATGT 1250

Qy   1251 TGCCCAGGCTGGTCTCGAAACTTCTGGCCTCATGCAATCCTCCCAAATGAGCCACACAACC 1306
          |||||||||||||||||||| ||||||||| |||||||||||||||||||||||||||||
Db   1247 TGCCCAGGCTGGTCTCAAACTTCTCGGCTCAGTCAGATCTCCCAAATGAGCCACCACAACC 1306

Qy   1307 CAGCCAGTCACATTTTTTTAAACAGTTACAATCTTTATTTTTAGTATTACTAGAAAGTAATACA 1370
          |||||||||||||||||| |||||||||||||||||||||||||||| |||||||||||||||
Db   1307 CAGCCAGTCACATTTTTTTAAACAGTTACAATCTTTTATTTTTAGTATTACTAGAAAGTAATACA 1366

Qy   1371 ATAAACATGTCAAAC 1385
          |||||||||||||||
Db   1367 ATAAACATGTCAAAC 1381
```

RESULT 6
US-08-340-426D-49
; Sequence 49, Application US/0834026D
; Patent No. 5948634
; GENERAL INFORMATION:
;   APPLICANT: de la Monte, Suzanne
;   APPLICANT: Wands, Jack R.
;   TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
;   TITLE OF INVENTION: of Alzheimer's Disease
;   NUMBER OF SEQUENCES: 121
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, Suite 600
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-3934
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/340,426D
;   FILING DATE: 14-NOV-1994
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Ludwig, Steven R.
;   REGISTRATION NUMBER: 36,203
;   REFERENCE/DOCKET NUMBER: 0609.3840002
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 49:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1381 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: both
US-08-340-426D-49

```
Query Match          74.98;   Score 1080.2;   DB 2;   Length 1381;
Best Local Similarity   94.4%;   Pred. No. 3.5e-292;
Matches 1317;  Conservative  0;  Mismatches  53;  Indels  25;  Gaps  18;

Qy        2 TTTTTTTTTGAGATGGAGAGTTTCGTCTTGTTGGCTGGAGTGCAAATGGCGCAAT 61
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 TTTTTTTTTGAGATGGAGAGTTTCGTCTTGTTGGCTGGAGTGCAAATGGCGCAAT 60

Qy       62 CTCAGCTCACCCGCCAACCTCCGGTCCTCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 121
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTCAGCTCACCCGCCAACCTCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy      122 AGTAGCTGGGATTACAGGCATGTGCCACCACGCCTGGCTAATTTTTGTATTTTTAGT 181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AGTAGCTGGGATTACAGGCATGTGCCA-CCACGCCTGGCTAATTTTTGTATTTTTAGT 179

Qy      182 AGAGATGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAACTCCCAGACCTCAGATGAT 239
            |||||||||||||    ||||||||||||||||||||||||||||||||||||||
Db      180 AGAGATGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCAGACCTCAGATGAT 239

Qy      240 CCCTCCGTCTCGGCCTCCCAAAGTGCT---AGATACAGGACTGCCACCATGCCCGG-CT 295
            |||||||||   |||||      |||       | |   |   |||||||||||| ||
Db      240 CTCCCGTCTCGGCCTGCCGAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCT 299

Qy      296 CTGGCCTGGCTAATTTTTGTGGTAGAGAAACAGGGTTTCACTGATG-TGCCCAAGCTGGTCTC 354
            ||||||||||||||||||||    || |||||||||||||||||| |||||||||| ||
Db      300 CTGGCCTGGCTAATTTTTGTGGTAGAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTC 359

Qy      355 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAG 414
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db      360 CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG 418

Qy      415 CCGGTGCCTGGCCTTTTATTTTTATTTTTTGAGACAGAGTCTCCACTCTTACCCAGG 474
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      419 CCGTGCCTGGCCTTTTATTTTTATTTTTTGAGACAGAGTGTACCACTCTTACCCAGG 478

Qy      475 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGATCAAGC-AT 533
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      479 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGATCAAGCAAT 538

Qy      534 CCTCCTGCCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTA 593
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      539 CCTCCTGCCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACACACACCACTACACCTGG-TA 597

Qy      594 ATTTTTTTATTTTTTTAGTTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGA 653
            |||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db      598 ATTTTTTTATTTTTTTAGTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA 656

Qy      654 GTGCAGTGGCGCAATCTCGGCTCACTGCAAGCTCTGCCTCCCGGGTTCAAGTGATTCTCC 713
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      657 GTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC 716

Qy      714 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCTAGCTAATTTTTTTT 773
            ||||||||||||||||||||||||||||||||||||||||||||||||||  ||||
Db      717 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCTAGCTAATTTTTTT 776

Qy      774 GTATTTTTAGTAGAGATGGGG-TTCACCATGTTGGCCAGGTTGATCTGGACC 832
            ||||||||||||||||||||| ||||||||||||||||||||||||||||
Db      777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATCTGGACC 836

Qy      833 TTGTGATCTGCCTGCGCCTCT-CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG 891
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      837 TTGTGATCTGCCTGCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCAC 894

Qy      892 CCCGGCTTATTTTTAATTTTTTGTTTGTTTGAAATGAAATGGAATCTCACTCTGTTACCCAGCGTCG 951
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      895 GCCGGCTTATTTTTAATTTTTTGTTTGTTTGAAATGAAATGGAATCTCACTCTGTTACCCAGCGTCG 954

Qy      952 GAGTGCAATGGCCAAATCTCGGCTCACTGCAACCTCCGGGCTCAAGCGATTCT 1011
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      955 GAGTGCAAATGG-CAAATTCTCGGCTCACTGCAACCTCTGCCTCCCGGG-TCAAGCGATTCT 1012

Qy     1012 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCGCACCACCACCCGCCACCACCCGCTAATT 1071
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1013 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCGCACCACCACCCGGGGCT-GGTTCTCAAGCTCCT 1070

Qy     1072 TTTGTATTTTTCATTAGAGGGCGGGGTTTCACCATGTTGGCCAGGCT-GGTCTCAAACTCCT 1130
            ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db     1071 TTTGTATTTTTCATTAGAGGGCGGG--TTTACCATAATTTGTCAGGCTGGTCTCAAACTCCT 1128

Qy     1131 GACCTCAGGTGAACCCACCTGCCTCAGCCTCTGCCTCCCAAAGTGCTGGATTACAGGCGTGAGCCA 1190
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1129 GACCTCAGGTGACCCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1188

Qy     1191 CCTCACCCAGCCGGGCTAATTTAGATAAAAAAAAAAATATGTAGCAATGGGGGGGTCTTGCTATGT 1250
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1189 CCTCACCCAGCCGGGCTAATTTGGAATAAAAAATATGTAGCAATGGGGG--TCTGCTATGT 1246

Qy     1251 TGCCCAGGCTGGTCTCAAAACTTCTGCCTTCAATGCAANCCTTCCAAATGAGGCCACAACGACC 1310
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1247 TGCCCAGGCTGGTCTCAAAACTTCTGGCTTCAGTCAATCCTTCCAAATGAGCCACAACACC 1306

Qy     1311 CAGCCAGTCACATTTTTTTAAACAGTTTACATCTTTATTTTAGTATATACTAGAAAGTAATTACA 1370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1307 CAGCCAGTCACATTTTTTTAAACAGTTACATCTTTATTTTAGTATATACTAGAAAGTAATTACA 1366

Qy     1371 ATAAACATGTCAAAC 1385
            |||||||||||||||
Db     1367 ATAAACATGTCAAAC 1381
```

RESULT 7
US-08-450-673C-49
; Sequence 49, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;   STREET: 1100 New York Avenue, Suite 600
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-3934
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/450,673C
;   FILING DATE: 30-MAY-1995
;   CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
;   NAME: Ludwig, Steven R.
;   REGISTRATION NUMBER: 36,203
;   REFERENCE/DOCKET NUMBER: 0609.3840004
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202) 371-2600
;   TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 49:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1381 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: both
;   TOPOLOGY: both
US-08-450-673C-49

Query Match   74.9%;   Score 1080.2;   DB 2;   Length 1381;
Best Local Similarity   94.4%;   Pred. No. 3.5e-292;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

```
Qy      2  TTTTTTTTTTTGAGATGGAGTAGTTTCGCTCTGTTGCCCAGGCTGGAGTGCAAATGGCGCAAT   61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTTTTTTTTTGAGATGGAGTAGTTTCGCTCTGTTGCCCAGGCTGGAGTGCAAATGGCGCAAT   60

Qy     62  CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTCAGCCTCCC          121
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTCAGCCTCCC          120

Qy    122  AGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTTAGT         181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTTAGT         179

Qy    182  AGAGATGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT     239
           |||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db    180  AGAGATGGAGTTTAACTCAGGCATGTGCA--CCACGCTCGGCTAATTTTTGTATTTTTAGT   239

Qy    240  CCCTCCGTCTCGGCGGCCTCCCAAAGTGCT---AGATACAGGACTGGCCACCATGCCCGG-CT   295
           ||||||||||||||||||||||||||||||    ||||||||||||||||||||||||| ||
Db    240  CTCCGCGTTCGGCGGCCTCCCAAAGTGCTGAGATTACAGGCATGAGCCACCATGCCCGACCT   299

Qy    296  CTGGCCTCAAGCTCACACTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCCGTGCAG      354
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300  CTGGCCTGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAGGCTGGTCTC   359

Qy    355  CTGAGCTCAAAGCAGTCACTGCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAG   414
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||  |||
Db    360  CTGAGCTCAAGCAGTCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGT-CAG        418

Qy    415  CCGTGCCTGGCCGTTTTTTAGTTATTTTAGTTATTTTTAAGACACAGGTGTCCACTCTTACCCAGG  474
           |||||||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419  CCGTGCCTGGCCTTTTATTTTTTAGTTATTTTTTAAGACACAGGTGTCACTCTTACCCAGG   478

Qy    475  ATGAAGTGCAGTGGTGTGATCACAGACGCTCACTGCAGCCTCGACATCAAGC-AT        533
           |||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db    479  ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCAACTCCTGAGATCAAGCAAT    538

Qy    534  CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACAGTCTCAACTCTGTGTCACCAGGCTA   593
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Db    539  CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGG-TA    597

Qy    594  ATTTTTATTTTTATTTTTATTTTTGAGACAGAGTCTCAACTCTGTCGTCCCAGGCTGGA     653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598  ATTTTTATTTTTATTTTTGAGACAGAGTCTC-ACTCTGTTCACCCAGGCTGGA          656

Qy    654  GTGCAGTGGCGCGAAATCACTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC   713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657  GTGCAGTGGCGCCAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCC     716

Qy    714  TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGCGCGCCACCACGCCCACCTAATTTTTTT     773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717  TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCCAGCTAATTTTTT      776

Qy    774  GTATTTTTAGTAGAGAGATGGGGTTTCACCATGTTGATTCTGATCTCTGGACC        832
           |||||||||||||||||||||||||||||||||||||||||||
Db    777  GTATTTTTAGTAGAGATGGGGTTTCGCCATGTTGATCTCTGATCTCTTGAACC          836

Qy    833  TTGTGATCTGCCTCGGCCTCT-CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG    891
           ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    837  TTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACACTGTGTTACCAGGCTG    894

Qy    892  CCCGGCTTATTTTTATTTTTGTTTTGTTTTGAGATGGAGTGTCGGGATTACAGGCTG     951
           ||||||||||||||||||||||||||||||||||||  |||||
Db    895  GCCGGCCTATTTTAATTTTTGTTTTGTTTTGAGAATGCTAAGCTCCTGTTTACCCAGGCTG   954

Qy    952  GAGTGCAAATGGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGGATTCT   1011
           |||||||||||||||||||||||||||||||||||||||||||
Db    955  GAGTGCAAATGG-CAAAATCCTGGCTACTGCAACCTCCGCCTCCCGGG-TCAAGCGATTCT    1012

Qy   1012  CCTGTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCGACCTGCACCACCACACCCGGCTAATT  1071
           ||||||||||||||||||||||||||||||||||||||||||
Db   1013  CCTGTCTCAGCCTCCCAAAGCAGCTGGGATTACGGGCGACCTGCACCACCACCCGGG--ACCTGCAC   1070

Qy   1072  TTTGTATTTTTCATTAGAGGCGGGGTTTCACCATATATTGTCAGGCT-GGTCTCAAAACTCCT   1130
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1071  TTTGTATTTTTCATTAGAGGCGGGG--TTTACCATATTGTCAGGCTGGTCTCAAAACTCCT    1128
```

PCT-US95-17111A-49

Query Match   74.9%; Score 1080.2; DB 5; Length 1381;
Best Local Similarity   94.4%;
Matches 1317; Conservative 0; Mismatches 53; Indels 25; Gaps 18;

```
Qy      2  TTTTTTTTTTTGAGATGGAGTAGTTTCGCTCTGTTGCCCAGGCTGGAGTGCAAATGGCGCAAT   61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTTTTTTTTTGAGATGGAGTAGTTTCGCTCTGTTGCCCAGGCTGGAGTGCAAATGGCGCAAT   60
```

```
Db    1 TTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCTCAGGCTGCAATGGCGAAT 60
Qy   62 CTCAGCTCACCGGCCACCTGCCTCACCGGTTCAAGGCAATTCTCCTGCCTCAGCCTCCC 121
Db   61 CTCAGCTCACCGCAACCTGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
Qy  122 AGTAGCTGGGATTACAGGCACATGTGCCACCACGCCCGGCTAATTTTGTAATTTTTTTAGT 181
Db  121 AGTAGCTGGGATTACAGGCGTGCA-CCACGCCCGGCTAATTTTGTATTTTTTTTTAGT 179
Qy  182 AGAGATGGGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAACCTCAGATGAT 239
Db  180 AGAGATGAGGGAGTTTTAACTCACTGTTGGTCAGGCTGGTCTCGAACCTCCCAGATGAT 239
Qy  240 CCCTCCGTCTCGGGCCTCCCAAAGTGCT---AGATACAGGACTGGCCACCACGCCCGG-CT 295
Db  240 CTCCGGTCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCT 299
Qy  296 CTGCCTGGCTAATTTTTTGTGGTAGAAACAGGGTTTCACTGAG-TGCCCAAGCTGGTCTC 354
Db  300 CTGCCTGGCTAATTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTC 359
Qy  355 CTGAGCTCAAAGAGAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAG 414
Db  360 CTGAGCTCAAAGCAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG 418
Qy  415 CCGTGGCCTCGGCGTCTTTTATTTTTATTTTTTTAAGCACACAGGTGTCCCACTCTTACCCAGG 474
Db  419 CCGTGGCCTGGCCTTTTTTATTTTTATTTTTTTTAAGACACAGGTGTACCACTCTTACCCAGG 478
Qy  475 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGACTCAAGC-AT 533
Db  479 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGATCAAGCAAT 538
Qy  534 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACCACCTGGCTA 593
Db  539 CCTCCCTGCCTCAGCCTCCCAAGTAGCTGGGACCCAAAGACATGCACCACTACCACCTGG-TA 597
Qy  594 ATTTTTTATTTTTTTATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGA 653
Db  598 ATTTTTTATTTTTTAATTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA 656
Qy  654 GTGCAGTGGCGCAATCTTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTCTCC 713
Db  657 GTGCAGTGGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC 716
Qy  714 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTTTT 773
Db  717 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTTTT 776
Qy  774 GTATTTTTAGTAGAGATGGGG--TTCACCATGTTGGCCAGGTTGATCTCTGGACC 832
Db  777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATCTTCTTTGACC 836
Qy  833 TTGTGATCTGCCTGCCTCGGCCT-CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG 891
Db  837 TTGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAG--GTCGTGACTCCAC 894
Qy  892 CCCGGCCTATTTTTTAATTTTTTGTTTTGAAATGATCTCACTCTGTTACCCAGCTG 951
Db  895 GCCGGCCTATTTTTAATTTTTTGTTTGTTTGAAATCTCACTCTGTTACCCAGGTCG 954
Qy  952 GAGTGCAATGGCGAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCT 1011
Db  955 GAGTGCAATGG-CAAATCTCGGCTACTGCAACCTCTGCCTCCCGGG-TCAAGCGATTCT 1012
Qy 1012 CCTGTGTCAGCCTCCCAAAGCAGCTGGGATTACAGGCGACCTGCCACCACACCCGCTAATT 1071
Db 1013 CCTGTGTCAGCCTCCCAAGCAGCTGGGATTACGGGGCACCTGCCACCACACCCGCTAATT 1070
Qy 1072 TTTGTAATTTTCATTAGGGACGGGGTTTCACCATATATTTGTGAGATAGAG-GGTCTCAAACTCCT 1130
Db 1071 TTTGTAATTTTCATTAGAGGCGGGG--TTTACCATATATTTGTCAGGCTGGTCTCAAACTCCT 1128
```

```
Qy 1131 GACCTCAGGTGACCCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1190
Db 1129 GACCTCAGGTGACCCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1188
Qy 1191 CCTCACCCAGCGCCGGCTAATTTTAGATAAAAAAAATATGTAGCAANTGGGGGGTCTTGCTATGT 1250
Db 1189 CCTCACCCAGCGCCGGCTAATTTTGGAATAAAAAATATGTAGCAANTGGGGG--TCTGCTATGT 1246
Qy 1251 TGCCCAGGCTGGTCTCGAAACTTCTCGGCCTTCATGCAAATCCTTCCAAANTGAGCGCACAACACC 1310
Db 1247 TGCCCAGGCTGGTCTCGAAACTTCTCGGCCTTCAGTCAANTCCTTCCAAATGAGCGCACAACACC 1306
Qy 1311 CAGCCAGTCACATTTTTTAAACAGTTACATCTTTATTTTAGTTATACTTAGGTAATACA 1370
Db 1307 CAGCCAGTCACATTTTTTAAACAGTTACATCTTTATTTTAGTTATACTTAGGTAATACA 1366
Qy 1371 ATAAACACATGTCAAAC 1385
Db 1367 ATAAACACATGTCAAAC 1381
```

```
Qy    1 TTTTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
Db 5380 TTTTTTTTTTTCTGAGATAGAG-TTTCACTCTGTTGTCCCAGGCTGGAGTGCAAATGGTGCAA 5438
```

```
Qy      61 TCTCAGCTCACCGCAACCTTCCGCCTCCCGGGTTCAACGCGATTCTCCTGCCTCAGCCGTCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5439 TCTTGGCTCACTGCAACCTTCCGCCTCCCGGGTTCAAGCGCGATTCTCCTGCCTCAGCCTCC 5498

Qy     121 CAGTAGCTGGGATTACAGGCGCATGTGACCACGCTCGGCTAATTTTGTATTTTTTTTTTAG 180
           |||||||||| ||||||||||| |||||||| ||||||| ||||||||||||||||||||||
Db    5499 AAGTAACTGGGATTACAGGCGCCCACCACCAGCGCCCAGCTAATTTTTTGTATTTTTTAG 5555

Qy     181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACCTCCGACCTCAGATGATC 240
           ||||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db    5556 TAGAGATGGGGTTTCACCACATTGGTCAGGCTGGTCTTGAACTCCTG---TGTAGTTTTAG 5613

Qy     241 CCTCCGGTCTCGGGCCTCCCAAAGTGCTAGGATACAGGACTGGGCCACCACGCCCGGCTCC 300
           |||||||| ||||| ||||||||||||||| ||||||||||| ||||||||||| |||| ||
Db    5614 CGCCCACCTTGGCCTCCCAAAGTGCTGG--GATTACAGGCGTGAACCACCGCTGGC 5670

Qy     301 TGGGCTAATTTTTGTGTAGAGAAACAGGGTTCACTGAGTTAGGCCAAGCTGGTCTCCTG--- 357
           ||| | ||||||||||||||||||||||||||||||||||||| | ||||||||||||
Db    5671 TTTTTTTTTTTTTTGTTCT-GAGACACAGTTTCACTCTGTTACCCAGGCTGGAGTAGGTGG 5729

Qy     358 -AGCTCAAGCAGTGCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCGAGC 415
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5730 CCTGATCTCGGATCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTTCAGC 5789

Qy     416 CGTGCCCTGGCCTTTTTATTTTTTTATTTTTTAAGACACAGGTGTCCCACTCTTACCAGGA 475
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5790 CTCCCAAGTAGCCGAGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTT 5849

Qy     476 TGAAGTGCAGTGGTGTGATCA--CAGCTCACTGCAGCCTTCAACTCCTGAGATCAAG-CA 532
           |||| | ||||||||||||||  ||||||||||||||||||||||||||||||||| ||
Db    5850 GGTAGACAGAGTTTCACCACATGTGGGCCAGGCTGGTTTGAACTCCTGACCTCAGTGA 5909

Qy     533 TCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCACAAAGACATGACCACCACTGGGC- 591
           ||| ||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    5910 CAAGCCAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGTGTGAGCCACCACA 5969

Qy     592 -------------TAATTTTTTATTTTTATTTTTTGAGACAGAGTCTCAACTCTGT 640
                        |||||||||||||||||||||||||||||||| |||||||||
Db    5970 TCAGGAAGTATTTTTTTTAAATTTTTATTTTTATTTTTTGAGATGGAGTCT-TGCTCTGT 6028

Qy     641 CACCCAGGCTGGAGTGCAGTGGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTT 700
           |||||||||||||||| |||||||||||||||||||| ||||||||||||||||| |||||
Db    6029 CGCCCAGGCTAGAGTGCAGCGACGGGAGTCTCGGCTCACTGCAAGCTCCGCCTCCCCAGGTT 6088

Qy     701 CAAGTGATTCTCCTGCCCGCCTCCTGAGTAGCTGGGACTACAGGCGC---CCACCACG 757
           ||| |||||||||||||||||||||||||||||||||||||||||||   |||||||
Db    6089 CAAGCCAATTCTCCTGCCTCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCGCCACCACA 6148

Qy     758 CCTAGCTAATTTTTGTATTTTTTAGTAGAGATGGG-GTTCACCATGTTCGCCAGGTTGA 816
           || | || ||||| ||||| ||| |||||||||| |||||||||||| ||||||||
Db    6149 CCCGGCTAATTTTTGTATTTTTTAGTAGAGACGGGGTTTCACCGTGTTTAGCCAGGAGG 6208

Qy     817 TCTTGATCTCCTGGACCTTGTGATCTCGCCTCGCCTCCCAAAGTGCTGGGATTACAG 876
           |||| |||||||| |||||||||||| |||||||||||||||||||||||||||||
Db    6209 TCTTGAACTCCTGACCTCGTGAATCCGCCCGCCTCCGGCCTCCCAAAGTGCTGGGATTACAG 6268

Qy     877 GCGTGAGCCACCACACGCCCGGCCTTATTTTTTAATTTTTGTTTGTTTGAAATGGAGTTACT 936
           |||| |||||||||||||||| ||| || | ||||||| |||||| |||||| |||||
Db    6269 GTGTGAGCCACCACACCCGGCT------ATTTT-TATTTGTTTTTTTGAGACGGGGTCTCACT 6320

Qy     937 CTGTTACCCAGGCTGGAGTGCAAATGGCACGATCTCGGCTCACTGCAACCTCTGCCTCCCG 996
           || ||| ||||||||| ||||||||||||| ||||||| ||| |||||||||||||||||
Db    6321 CTGTCACCCAGGCTGCAGTGG-TACACCATAGCTCACTGCAGCCTCGAACTCCTG 6379

Qy     997 GGCTCAAGCGATTCTCCTGTCTCGGCCTCCCAAGCAGCTGGGATTACGGGCACTGCCAC 1056
           |||||||||||||||| ||||||||||||||||| |||||||||||| ||||||||||
Db    6380 AGCTCAAGTGATCCTCCCACCTCAGTAATTGGGACTACAGGTGCACCCAC 6439

Qy    1057 CACACCCCCGGCTAATTTTTTATTTTTTGTAT--------------TTTCATTTAGAGGCG 1101
           ||| |||| |||||||||| | | ||| |||                |||| |||||||| 
Db    6440 CATGCCCACCTAATTTTTATTTTTATTTTCATAGAGACGGGGTTTCAC 6499

Qy    1102 CATATTTGTCAGGCTGGTCTCAAAACTCCTGACCTCAGGTGACCACCTGCCTCAGCCTTC 1161
           |||| |||| |||||||| || |||| |||| ||||||||| ||||| ||||||||||| 
Db    6500 TGTGTGTGTCCAGGCTGGTGTCTTGAAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTCC 6559

Qy    1162 CAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACCTCACCCAGCCGGTCAAATGCGCCA 1218
           |||||||||||||||||||||| ||| ||||||||||| || |||| |||| |||||||
Db    6560 CAAAGTGCTGAGATTACAGGCCATGAGCCACCGCACCTAGGAATCATTTTTAAA 6616
```

RESULT 10
US-09-630-706-10
; Sequence 10, Application US/09630706
; Patent No. 6277640
; GENERAL INFORMATION:
;   APPLICANT: C. Frank Bennett
;   APPLICANT: Lex M. Cowsert
;   TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION
;   FILE REFERENCE: RTS-0053
;   CURRENT APPLICATION NUMBER: US/09/630,706
;   CURRENT FILING DATE: 2000-08-01
;   NUMBER OF SEQ ID NOS: 94
; SEQ ID NO 10
;   LENGTH: 14796
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (2811)...(2921)
;   NAME/KEY: CDS
;   LOCATION: (3174)...(3283)
;   NAME/KEY: CDS
;   LOCATION: (5158)...(5275)
;   NAME/KEY: CDS
;   LOCATION: (11955)...(12044)
US-09-630-706-10

Query Match   33.0%;   Score 475.4;   DB 4;   Length 14796;
Best Local Similarity   69.8%;   Pred. No. 7.8e-123;
Matches 877;   Conservative 0;   Mismatches 321;   Indels 59;   Gaps 15;

```
Qy       1 TTTTTTTTTTTTGAGATGGGAGTTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAAATGGCGCA 60
           ||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||
Db    5380 TTTTTTTTTTTTCTGAGATAGAG-TTTCACTCTTTGTTGCCCAGGCTGGAGTGCAATGGTGCAA 5438

Qy      61 TCTCAGCTCACCGCAACCTTCCGCCTCCCGGGTTCAACGCGATTCTCCTGCCTCAGCCGTCC 120
           ||| ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    5439 TCTTGGCTCACTGCAACCTTCCGCCTCCCGGGTTCAAGCGCGATTCTCCTGCCTCAGCCTCC 5498

Qy     121 CAGTAGCTGGGATTACAGGCGCATGTGACCACGCTCGGCTAATTTTGTATTTTTTTTTTAG 180
           |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    5499 AAGTAACTGGGATTACAGGCGCCCACCACCAGCGCCCAGCTAATTTTTTGTATTTTTTAG 5555

Qy     181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACCTCCGACCTCAGATGATC 240
           ||||||||||||||||| ||||| |||| ||||||||||||||||||||||||||||||
Db    5556 TAGAGATGGGGTTTCACCACATTGGTCAGGCTGGTCTTGAACTCCTG---TGTAGTTTTAG 5613

Qy     241 CCTCCGGTCTCGGGCCTCCCAAAGTGCTAGGATACAGGACTGGGCCACCACGCCCGGCTCC 300
           ||||||| |||||||||||||||||||| ||||||||||||||| |||||||||||| || ||
Db    5614 CGCCCACCTTGGCCTCCCAAAGTGCTGG--GATTACAGGCGTGAACCACCGCTGGC 5670

Qy     301 TGGGCTAATTTTTGTGTAGAGAAACAGGGTTCACTGAGTTAGGCCAAGCTGGTCTCCTGG--- 357
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5671 TTTTTTTTTTTTTTGTTCT-GAGACACAGTTTCACTCTGTTACCCAGGCTGGAGTAGGTGG 5729

Qy     358 --AGCTCAAGCAGTGCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCGAGC 415
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5730 CCTGATCTCGGATCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTTCAGC 5789

Qy     416 CGTGCCCTGGCCTTTTTATTTTTTTATTTTTTAAGACACAGGTGTCCCACTCTTACCAGGA 475
           ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5790 CTCCCAAGTAGCCGAGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTT 5849

Qy     476 TGAAGTGCAGTGGTGTGATCA--CAGCTCACTGCAGCCTTCAACTCCTGAGATCAAG-CA 532
           |||| | ||||||||||||||  ||||||||||||||||||||||||||||||| ||
Db    5790 CTCCCAAGTAGCCGAGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTT 5849
```

RESULT 11
US-09-496-694B-3
Sequence 3, Application US/09496694B
Patent No. 6335194
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Elizabeth J. Ackermann
APPLICANT: Eric E. Swayze
APPLICANT: Lex M. Cowsert
TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION
FILE REFERENCE: ISPH-0439
CURRENT APPLICATION NUMBER: US/09/496,694B
CURRENT FILING DATE: 2000-02-02
PRIOR APPLICATION NUMBER: 09/286,407
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 09/163,162
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 249
SEQ ID NO 3
LENGTH: 14796
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: CDS
LOCATION: (2811)...(2921)
NAME/KEY: CDS
LOCATION: (3174)...(3283)
NAME/KEY: CDS
LOCATION: (5158)...(5275)
NAME/KEY: CDS
LOCATION: (11955)...(12044)
US-09-496-694B-3

Query Match 33.0%;   Score 475.4;   DB 4;   Length 14796;
Best Local Similarity 69.8%;   Pred. No. 7.8e-123;
Matches 877;   Conservative 0;   Mismatches 321;   Indels 59;   Gaps 15;

[Sequence alignment data (Qy = Query, Db = Database) follows]

RESULT 12
US-09-813-817-3/c
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match   28.9%;
Best Local Similarity   67.7%;
Matches 815; Conservative 0; Mismatches 326; Indels 63; Gaps 14;
Score 416.4; DB 4; Length 59065;
Pred. No. 4.3e-106;

```
Qy      42 CTGGAGTGCAATGGCGCAAATCTCAGCTCACCCGAACCTCCCGGGTTCAAGCGAT 101
           |||||| |||| |||||| ||||||||||||||||||||| || |||| |||||||
Db   15163 CTGGAGGGCAGTGGCGCGATCTCAGCACACTCCATCTCCCAGGTTCAAGCTAT 15104

Qy     102 TCTCCTGCCTCCAGCCTCCCAGTAGCTGGGATTACAGGCACGCACCGCCGGCTA 161
           || ||| ||||||||||||| |||||||||||||| ||||| ||||||||||||
Db   15103 .TCTCCTGCCTTAGTCTCTCAAGTAGCTGGGATTACAGGTGCCTGCCA 15044

Qy     162 ATTTTGTATTTTTTTTAGTAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAA 221
           |||| | |||||||||||| ||| ||||| ||||||||||||||||| |||
Db   15043 ATTTT---TGTAATTTAGTAGRAGACAAGGTTTCACCATGTTGGCCAGGCTGGTCTGGAA 14987

Qy     222 CTTCCCGACCTCAGATGATCCCTCGGCCTCCCAAAGTGCT-AGATACAGGACTGG 280
           || ||||||||||| |||| |||||| |||||||| ||| | |||||||||||| |
Db   14986 CTCCTGACCTCAGGTGATCCCCAAAGTGCTGGGATTACAGGCGCG 14927

Qy     281 CCACCATGCCCGGCCGTCGTGCCTCGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTG 340
           || || ||| |||| || |||||| ||||||| ||| ||| ||| ||||||||||||
Db   14926 AGCCACCGCGGCCCAACCTCCACATTTTTTTTTTTTTTTGAGACAGAGTCTTGCTCTGTCG 14867

Qy     341 CCCAAGCTG------GTCTCCTGAGCTCA 363
           |||| ||||      | ||| | |||
Db   14926 AGCCACCGCGGCCCAACCTCCAC 14867
```

RESULT 13
US-09-978-197-3/c
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178DIV

CURRENT APPLICATION NUMBER: US/09/978,197
CURRENT FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 09/813,817
PRIOR FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 3
LENGTH: 59065
TYPE: DNA
ORGANISM: Human
US-09-978-197-3

Query Match 28.9%; Score 416.4; DB 4; Length 59065;
Best Local Similarity 67.7%; Pred. No. 4.3e-106;
Matches 815; Conservative 0; Mismatches 326; Indels 63; Gaps 14;

Qy   42 CTGGAGTGCAATGGCGCAAATCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGAT   101
          |||| || || || |||||||||||||||||||||||| |||||| ||||||||||||||
Db 15163 CTGGAGGGCAGTGGCGCCGACTCAGCACACTGCAACTCCAATCTCCCAGGTTCAAGCTAT 15104

Qy   102 TCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCATGTGACCACCACGCTCGGCTA   161
          |||||||||| |||||| |||||| |||||||||||||| |||||| |||||| |||||
Db 15103 TCTCCTGCCTTAGTCTCTCAAGTAGCTGGGATTACAGGTGCCTGCCATCGTGCCCGGCTA 15044

Qy   162 ATTTTGTATTTTTTTAGTAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAA   221
          ||||| ||||||||| ||||||||||| | |||| ||||||||||||||||||||| ||
Db 15043 ATTTT---TGTAATTTTAGTAGAGACAAGGTTTCACCATGTTGGCCAGGCTGGTCTCGGAA 14987

Qy   222 CTCCCGACCTCAGATGATCCCTCCGTGTCGGCCTCCCAAAGTGCT-AGATACAGACTGG   280
          |||||| ||||||||| ||| ||| || ||||||||||||| | ||||||| |||||
Db 14986 CTCCTGACCTCAGGTGATCCGCCCGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCGCG 14927

Qy   281 CCACCATGCCCGGCTCTGCCTGCTAATTTTTGTGTAGAAACAGGGTTTCACTGATGTG   340
          ||| || |||||||| ||||||||||||||||||||| |||||||||||||||| |||
Db 14926 AGCCACCGGCGCCCAACCTCCATTTTTTTTTTTTTGAGACGAGTCTTGCTCTGTCG 14867

Qy   341 CCCAAGCTG-----------------------GTCTCCTGAGCTCA   363
          |||||||| || | |||||||||
Db 14866 CCCAGGCTGGACTGCAGTGGCGCGATCTCAGCTCACTGCAAGCTCA 14807

Qy   364 AGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAG--GCGTGCAGCCGTGCC   421
          || |||| || |||||||||||||||||||||||||||||||| | ||| |||| ||
Db 14806 CGCCATTCTCCTGCCTCAGCCTCGAGTAGCAGGGACTACAGGCGCCCGCCACCACATCC 14747

Qy   422 TGGCCTTTTTAGTTATTTTTTTTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGAAGT   481
          ||||| |||| |||||||||||||||| | ||||||| | | || |||||||| || |
Db 14746 TGGCCTAATTTTGTGTATTTTTTAGTAGAGACGGGGTTTCACTGTGTTAGCCAGGATGGTCT 14687

Qy   482 GCAGTGGTGTGATCACAGCTCACTGCACCTCAAACTCCTGAGATCAAGCATCCTCCTGC   541
          | || || | | |||| |||||||| || | ||||| |||||| |||| ||||||||
Db 14686 -CGATCTCCTGACCTTGTGATCCCGCCGCCTCGGCCTCCCAAAGTGCTGGGATTACATGC 14628

Qy   542 CTCAGCCTCCCAAGTAGCTGGGACCAAAGA-----CATGCACCACTACACCTGGCTAATT   596
          |||||||||||| |||||||||| ||||| | | || |||||||||||||||
Db 14627 GTGAGCCACCGCGTGCTGGCCTCCAATTTTTTGAGACGGAGTCTCGCTCTGTCGCCCAGG 14568

Qy   597 TTTATTTTTTATTTTTTTTAATTTTTTGAGACAGAGTCTCAACTCGTCACCCAGGCTGGAGTG   656
          ||||||||||||||||||| ||||||||||||| ||||| | | |||||| |||||||
Db 14567 TATTTTTAGTGTTATTTATTTATTTATTGAGACGAGTCTC-GCTCTGTCTCCCAGGCTGGAGTG 14509

Qy   657 CAGTGGCGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCCTGC   716
          || |||||| | |||||||||||| ||| | |||||| |||||| |||||||||||
Db 14508 CAGTGACAATGTTGGCCCACTCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGC 14449

Qy   717 CCCAGCCTCC----TGAGTAGCTGGGACTACAGGCGC----CCACCACGCCTGGCTAATTTTT   770
          ||||||||| ||||||||||| ||||||||||| ||| ||||| |||||||||
Db 14448 CTCAGCCTCCAAGTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCAGCTAATTTTT 14391

Qy   771 TTTGTATTTTTTAGTAGAGATGGGG--TTCACCATGTTGGTCAGGCTGGTCTCGAACTCCTG   829
          |||| ||||| ||||||||||||| |||||||||| |||| || |||||| |||||||||
Db 14390 TTTGTATTTTCAGTTGAGATGGGGTTTCACCATGTTGGCCAGGCTGCTGATCTCGAACTCCTG 14331

Qy   830 ACCT--TGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC   887
          |||| ||||||||||| |||| || ||||||||||| ||| ||||||| |||||||||
Db 14330 ACCTCAAGTGATCATCACCTGCCTTGGCCTCCCAAAGTGCTGGGATTACTGGCGTGAGCCAA 14271

Qy   888 CACGCCCGGCCTTATTT-TTAATTTCGTTTCTTTGTTGTTTTGAAATGGAAATCTCACTCTGTTACCCA   946
          ||||||| ||| ||| ||||||| || | |||||| | ||||||| | | | ||||| |||||||
Db 14270 CGCGCCCTGGCTCGTTCTTCTTTTTTTTTTTTTTTTTTTTGAGATAGGGTCTTGCTCTGTCATCCA 14211

Qy   1007 ATTCTCCTTGTCTCAGCCTCCAAGCAGCTGGGATTACGGGCACCTGCCACCACACCCGGC   1066
          || |||| ||||| || |||| |||||||||||||| |||||| | |||||| |||||
Db 14150 ATCCTCCTCGGCCTCCCAGCCTCGGAGTAGCTGGGACCACAGGCGCATGAGCCACCACGGC 14091

Qy   1067 TAATTTTTTGTATTTCATTAGAGACGGCGGGGTGTTCACCATATTTGTCAGGCTGGTC-TCAA   1125
          |||||||| | ||||| || ||||||| ||||||||||| || ||||| || ||| |||
Db 14090 TAAATTTTTTGATTTTTTTGTGGGGGATGAGGGTCTTGCTCTGTTACCCAGACTGGTCGGTTGAAA 14031

Qy   1126 CTCCTGACCTCAGGTGACCCACCTCAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCA   1185
          ||| ||||||||| |||| || ||| ||||||||| || ||||||||||| ||| ||||
Db 14030 CTCCTGAGTTCAAGCAAATCTCCCGGCCTCCCAAAGTGCTGGGATTATATAAGTGTA 13971

Qy   1186 AGCC   1189
          ||||
Db 13970 AGCC 13967

RESULT 14
US-09-735-934A-3/c
Sequence 3, Application US/0973593494A
Patent No. 6372468
GENERAL INFORMATION:
APPLICANT: LI, Jiayin et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL000851
CURRENT APPLICATION NUMBER: US/09/735,934A
CURRENT FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 3
LENGTH: 43950
TYPE: DNA
ORGANISM: Homo sapiens
US-09-735-934A-3

Query Match 28.68%; Score 412.2; DB 4; Length 43950;
Best Local Similarity 66.0%; Pred. No. 5.6e-105;
Matches 863; Conservative 0; Mismatches 333; Indels 111; Gaps 14;

Qy   2 TTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCCCAAT   61
          ||||||||||| ||||||||||| || |||||| ||||||||||||||||||| |||||
Db 37294 TTATTATTTTTTTTGTGGGTGTCGTCTGGTCTCGTTGCGGAGTTGCAGTGGTGTGTAT 37235

Qy   62 CTCAGTCACCCAACCTCCG-CCTCCCGGGTTCAAGGGATTCTCACTCAGAGTAGCTCCC   120
          |||||| |||| ||| | |||||||||||| || ||||| || | | || || |
Db 37234 CTCGGCTCACCGCAACCTCCACCTCCCGGGTTCAAGCGATTCCCCGCTTCAGCCTCCC 37175

Qy   121 CAGTAGCTGGGATTACAGGCATGTGCACCCACGCTCGGCTAATTTTTGTATTTTTTAG   180
          |||||||| |||||| |||| ||| |||||| |||||||||||||| ||||| |||
Db 37174 GAGAAGCTGGGATTACAGGTTCCCGCACCATGCCCGGCTAATTTTTGTATTTTTTTAG 37118

Qy   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATC   240
          ||||| ||||||| || |||| ||| | || |||||| |||||| | || |||||||
Db 37117 TAGAGACGGGGTTTCGGCCATGTTAGCCAGGCTGGTCTCCAACTCCCTGACCTCAAGTGATC 37058

Qy   241 CCTCCGCTCGGCCTCCCAAAGTGCTAGATTACAGGACTGGGCACCACCATGCCCGGCTGCC   300
          |||| | |||||||||||||||| || ||||| | || | ||||||| ||| ||| ||
Db 37057 TGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTA----TAGGCATGAGCCGCTGCGCCCAA 37002

Qy   301 TGGCTAATTTTTGTGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCCTGAGC   360
          |||||||||||| | || |||||||||||||| ||| |||||||||||||| | | ||||
Db 37001

~ APPLICANT: Gregory Dolganov
~ TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
~ NUMBER OF SEQUENCES: 175
~ CORRESPONDENCE ADDRESS:
~   ADDRESSEE: Dehlinger & Associates
~   STREET: 350 Cambridge Avenue, Suite 250
~   CITY: Palo Alto
~   STATE: CA
~   COUNTRY: USA
~   ZIP: 94306
~ COMPUTER READABLE FORM:
~   MEDIUM TYPE: Floppy disk
~   COMPUTER: IBM PC compatible
~   OPERATING SYSTEM: PC-DOS/MS-DOS
~   SOFTWARE: Patentin Release #1.0, Version #1.25
~ CURRENT APPLICATION DATA:
~   APPLICATION NUMBER: US/08/687,080
~   FILING DATE: 17-JUL-1996
~   CLASSIFICATION: 435
~ PRIOR APPLICATION DATA:
~   APPLICATION NUMBER: US 08/592,126
~   FILING DATE: 26-JAN-1996
~ ATTORNEY/AGENT INFORMATION:
~   NAME: Sholtz, Charles K.
~   REGISTRATION NUMBER: 38,615
~   REFERENCE/DOCKET NUMBER: 4600-0111.30
~ TELECOMMUNICATION INFORMATION:
~   TELEPHONE: (415) 324-0880
~   TELEFAX: (415) 324-0960
~ INFORMATION FOR SEQ ID NO: 101:
~   SEQUENCE CHARACTERISTICS:
~   LENGTH: 5543 base pairs
~   TYPE: nucleic acid
~   STRANDEDNESS: double
~   TOPOLOGY: linear
~   MOLECULE TYPE: DNA (genomic)
~   HYPOTHETICAL: NO
~   ANTI-SENSE: NO
~   ORIGINAL SOURCE:
~   INDIVIDUAL ISOLATE: 3' END OF INTRON 19 OF RAD50 GENOMIC
~   INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-101

Query Match 27.8%; Score 400.8; DB 2; Length 5543;
Best Local Similarity 72.1%;
Matches 610; Conservative 0; Mismatches 217; Indels 19; Gaps 6;

```
Db  37001 CCAGAAAACTTTTAAATATGTTGGAACAACTGTCAATCTATTTCCCAAACTG 36942
Qy    361 TCAAGCAGTCACCTGCCTCTCAGCCTCCCAAAGTGCCTGGGATTACAGGCGTGCAGCCGTGC 420
Db  36941 TTATACACACCAGATTTG------------AAGACATAGTATGAAAAACAAGACAGAAATGTT 36892
Qy    421 CTGGCCTTTTTAATTTTAATTTTTTTTTTAAGACAAGGTGTCCCACTCTTACCCAGGATGAAG 480
Db  36891 TCATTGAGTTTGTTTTCTTTTTTTTAAGACAGA-GGGTCTTGCTCTGTCTCACCCAGCTGGAG 36833
Qy    481 TGCAGTCGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAG-CATCCTCCT 539
Db  36832 TGCAGTGGCACAATCATAGCTCACTTCAGACTTGTAAACTCCTGGTCTCAAGCCGTCTTCCC 36773
Qy    540 GCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGC--------- 591
Db  36772 ATCTTGGCCTCCCAAAGTGCGGGGATTATAGGCATGAGCCACCTCCCCGGCCTTAATAA 36713
Qy    592 ------------------------------------------------------------ 591
Db  36712 TTTTTATGTTGACTACATGTTGAATAAAGTGCTATATGTTGGGATAAACTATATATATAT 36653
Qy    592 TAATTTTTTATTATTTTTTAATTTTTTTGAGACAGTCTC------AACTCTGTCACC 644
Db  36652 ATATATATATATATTTTTTTTTTTTTTTTGAGACAGAGTCTCGCTCTGTCTCGTCACC 36593
Qy    645 CAGGCTGGAGTGCAGTGGCGCGAATCTTGGCTCACTGCAAACCTCTGCCTCCCGGGTTCAAG 704
Db  36592 CAGGCTGGAGTGCAGTGGCGTGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCATG 36533
Qy    705 TTATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGC---CCACCAGCGCCC 761
Db  36532 CCATTCTCCTCAACTCACGCCTCCCAAGTAGCTGGGACTACAGGCGTGCGCCACCACGCCC 36473
Qy    762 GCTAA--TTTTTTTTTGTATTTTTTAGTAGAGAGGGG--TTCACATGTTCGGCCAGGTTGATC 818
Db  36472 GCTAATTTTTTTTTTTTGTTATTTTTTAGTAGAGAAGGGTTTCACCGTGTTAGCCAGGATGGTC 36413
Qy    819 TTGATCTCTCGGACCTTGTGATCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 878
Db  36412 TCGATCTCCTGACCTCGTGAACTGCCGCGCTCGGCCTCCCAAAGTGCTGGGATTACAGGC 36353
Qy    879 GTGAGCGCCACCACGCCCGGCCTTATTTTTAATTTTGTTTGTTTGAAATGGAATCTCACTCT 938
Db  36352 GTGAGCCACCGCACCC-----AGTCTTTTTTTTTTTTTTTAAGAGACAGCTCTCACTCT 36298
Qy    939 GTTACCCAGGCTGGAGTGCAAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGG 998
Db  36297 GTTATTCAGGCTGAGTGGCACAATCATAGCTCACTGCAGCCACCAACTCCTGG 36238
Qy    999 CTCAAGCGAGTTCTCCTGTCTCAGCCTCCCAAGCAGCAGGCACCTGCCACCA 1058
Db  36237 CTCAAGCAAATCTCCCACGTCTCCAAGCAGGCTAGGACTATAGGCGCGCACCA 36178
Qy   1059 CACCCGGCTAATTTTTGT---ATTTTCATTAGAGGCGGGGTTTCACCATATTTGTCAGGC 1115
Db  36177 TTCCCGGCTATTTTTTTATTTTTTGTATAGATGAGGTCTTGCCATGTTGCCCAGGC 36118
Qy   1116 TGGTCTCAAACTCCTGACCTCGTGATCAGCCTCCCAAAGTGCTGGGAT 1175
Db  36117 TGGTCTCAAACTCTTGGGCTCAAATGATCCCTCCTCGCCTCTCAAAATGCT--GAT 36060
Qy   1176 TACAGGCGCGTGAGCCACCACTGTGCCTCGGCCCCACCATCTTGATTAGTATA 1222
Db  36059 TACAGGCAAGAGCCACCACTGTGCCTTGGCCCCCACCATCTTGATTAGTATA 36013
```

Qy   592 TAATTTTTTATTTTTTATTTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTG 651
Db   1274 TGACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCT-GGCTCTGTCACCCAGGCTG 1216
Qy   652 GAGTGCAGTGGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTTTCT 711
Db   1215 GAGTGCAGTGGTGCGATCTCGGCTCACTGCAAGCTCTGCCTCCCAGGTTCACGCCATTCT 1156
Qy   712 CCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCAC---GCCTAGCTAATT 768
Db   1155 CCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACCAGCCCTGGCTAATT 1096
Qy   769 TTTTTTGTATTTTGTTTAGTAGAGATGGGG-TTCACCATGTTGGCGATGGTCTCGATCTCT 827
Db   1095 TTTTGTATTTTTTAGTAGAGACGGGGGTTTCGGCGTGTTAGCCAGGATGGTCTCGATCTCC 1036
Qy   828 GGACCTTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 887
Db   1035 TGACCTCGTGACTCGCCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 976
Qy   888 CACGCCCGGCCTAT--------TTTTAATTTTGTTTGTTTGAAATGGAATCTCACTC 937
Db   975 CACGCCCTGGCCTGTTTTTGTTTCTGTTTTGTTTGTTTGTTTGAAGACGGAGTTTCACTC 916
Qy   938 -TGTTACCCAGGCTGGAGTGCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCG 996
Db   938 CACTGCCACCGCGCCCGGCTAGTGGGATTACAGGCGTGAGCCAC

```
Db   915 TTGTCACCCAGGCTGAAGTGCAATGGTGTGAATCTCGGCTCACTGCAAATCTTCTGCCTCCCA  856
             ||||||      |||||||   ||| ||||| ||||  ||||  |||||||||||||||||
Qy   997 GGCTCAAGCGAATTCTCCTGTCTCAGCCTTCCAAGCAGCTGGGATTACGGGCACCTGCCAC 1056
         ||  |||||||||||||||||||||||||||||||||||||||||||| | ||| |||
Db   855 GGTTCAAGCGATTTCTCCTGCCTCAGCCTCTGAGTACCTGGGATTACAGGCGCGTGTCAC  796

Qy  1057 CACACCCCGCTAATTTTTGTATTTTCATTAGAGGCGGGGTTTCACCATATTTGTCAGGCT 1116
         ||||| |      ||||||||||||||||| ||||||||||||||||||||||||| |||
Db   795 CACACCCTGGCTAATTTTTCTATTTTCAGTAGAGATGGGGTTTTACCATATTGGCCAGGCT  736

Qy  1117 GGTCTCAAACTCCTGACCTCAGGTGACCCACCTGCCTCAGCCTTCAAAGTGCTGGGATT 1176
         ||||    |||||||||||||||||||| ||||||||||||||| |||||||||||||||
Db   735 AGTCTTGAACTCCTGACCTCAGGTGATCCGTCCTGCCTTGGGCCTCCCAAAGTGCTGGGATT  676

Qy  1177 ACAGGGCGTGAGCCACCTCACCCAGCCGGGCTAATTTTAGATAAAAAAAATATGTAGCAATGGG 1236
         |||||| |||||| |||||||| ||   |||  ||||| || |||||||| ||| |||||
Db   675 ACAGGCATGAGTCACTGCGCCCGGCCTCCTCTCTTTATTGACTACTAGAATCTTCAGCA  616

Qy  1237 GGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAAACTTCTGGCTTCATGCAAATCCTTCCAAA 1296
            |   |   | ||| |||  ||||||||||| |||  |||||||||||||||||||||
Db   615 AGCATATCAGACTTCATGCAATACTTTTTATACACTTCT---CTCCTGTTTCATTACTTT  559

Qy  1297 TGAGCCACACACCCAGCCAGTCACATTTTTTAAACAGTTACATCTTTATTTTAGTATAC 1356
         |||   |  | | | ||  ||||||||||||||||||||||| |||| | | ||
Db   558 CTTGCCCTTATTTCTACACTGCCTGTGTTTCCCATTAATTTGAAATACATTTATCTTTGC  499

Qy  1357 TAGAAAGTAATACAATAAACATGTCAAACCTTCAGTAGTAACGAGAGTTCTTTTTA 1416
         || | |     ||| |  |||  |||||| |||| || |||||| ||| |||
Db   498 TCTATTCTTATTATAACTAAGTAAATAATTTCTGGAACAAGAAGGTTACAAAGTAAAACTAA  439

Qy  1417 TAACTT 1422
         || | |
Db   438 TACCAT 433
```

Run on: February 1, 2003, 20:03:23 ; Search time 330 seconds

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model
(without alignments)
9840.550 Million cell updates/sec

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 ttttttttttgagatggag.........ttaaacaaagcttaagca 1442

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   N_Geneseq_101002:*
    1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
    2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
    3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
    4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
    5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
    6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
    7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
    8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
    9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
   10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
   11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
   12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
   13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
   14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
   15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
   16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
   17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
   18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
   19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
   20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
   21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
   22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
   23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
   24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

SUMMARIES

Pred	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1442	100.0	1442	24	ABN89470	Neural thread prot
	2	1223.4	84.8	1418	17	AAT27738	Neural thread prot
	3	1078.6	74.8	1381	15	AAQ77883	Neural thread prot
	4	1077	74.7	1381	17	AAT27765	AD 10-7 human neur
c	5	587.8	40.8	6508	24	ABL62910	Breast cancer rela
c	6	587.8	40.8	6508	24	ABL64414	Stomach cancer rel
c	7	587.8	40.8	6508	24	ABL67668	Oesophagus cancer
c	8	524.4	36.4	6944	24	ABL66947	Lung cancer relate
c	9	524.4	36.4	6944	24	ABL68262	Kidney cancer rela
c	10	517.2	35.9	33747	22	AAK69279	Human immune/haema
	11	517.2	35.9	33747	22	AAK73093	Human immune/haema
	12	513.4	35.6	43545	24	ABK85018	Human cadherin-11k
	13	513	35.6	21470	23	ABK42270	Genomic sequence #
	14	511.6	35.5	5262	22	AAK71768	Human immune/haema
	15	511.6	35.5	5262	22	AAK71769	Human immune/haema
c	16	490.4	34.0	99014	24	ABN96931	Gene #3429 used to
	17	489.6	34.0	24167	22	ABA16032	Human nervous syst
	18	479.6	33.3	44820	24	AAS19703	Reference sequence
	19	479.2	33.2	51719	24	AAD31365	52Kb gene fragment
	20	479.2	33.2	51853	24	AAD31364	92Kb gene fragment
c	21	478.2	33.2	31853	22	AAI98993	Human excretory re
c	22	477.4	33.1	11319	24	AAK00025	Human digestive sy
c	23	477.4	33.1	11319	22	AAV27941	Survivin gene. Ho
c	24	477.4	33.0	11796	19	AAS21523	DNA encoding human
c	25	475.4	33.0	14796	22	AAH47531	Human Her-3 genomi
	26	475.4	33.0	14796	24	ABN96924	Gene #3422 used to
	27	475.4	33.0	14796	24	ABL65663	Lung cancer relate
	28	475.4	33.0	14796	22	ABL63326	Lung cancer relate
	29	475.4	33.0	14796	24	ABL68655	Kidney cancer rela
	30	475.4	33.0	14796	24	AAT22317	Human transporter
	31	475.4	33.0	26557	24	AAS26800	Human genomic DNA
c	32	471.2	32.7	12542	22	AAS26800	Human cardiovascul
c	33	471.2	32.7	15041	22	AAS35921	Rhesus gene locus:
	34	467	32.4	9236	22	AAS03689	Rhesus gene locus:
	35	466.8	32.4	9236	22	AAS03690	Rhesus gene locus:
	36	466.8	32.4	9236	22	AAS03688	Rhesus gene locus:
	37	467	32.4	15041	24	AAK67211	Human death-associ
	38	461.8	32.0	12026	22	AAS00624	Human immune/haema
	39	461.8	32.0	36221	22	AAK69566	Human immune/haema
	40	460.8	32.0	2368	22	AAK89230	Human digestive sy
	41	459.4	31.9	30110	22	AAD02697	Human glycosyl sul
	42	457.8	31.7	160552	22	ABK84756	Human cDNA differe
	43	457.6	31.7	81800	24	AAK81036	Human immune/haema
	44	457.4	31.7	38771	22	AAF21086	Human low adenosin
c	45	456.4	31.7	16310	21		Human

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
ABN89470
ID   ABN89470 standard; DNA; 1442 BP.
XX
AC   ABN89470;
XX
DT   02-SEP-2002 (first entry)
XX
DE   Neural thread protein encoding nucleotide sequence.
XX
PN   WO200234915-A2.
XX
PD   02-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US42813.
XX
PR   27-OCT-2000; 2000US-0697590.
XX
PA   (NYMO-) NYMOX PHARM CORP.
XX
PI   Fitzpatrick J, Averback P, Focht MSS, Bibiano R;

DE   Neural thread protein (NTP; Harlil peptide; Alzheimer's disease;
KW   Down's syndrome; neuroectodermal tumour; astrocytoma; glioblastoma;
KW   hypoxia; ischaemia; cerebral infarction; gene; ds.

FH   Key   Location/Qualifiers
FT   CDS   15..1142
FT   /*tag= a
FT   /product= "Neural thread protein"
XX
OS   Homo sapiens.

```
XX
DR   WPI; 2002-507998/54.
DR   P-PSDB; ABB81538.
XX
PT   New Harl1l peptide sequences of the Neural Thread Protein, useful in
PT   therapeutic assays, e.g. as targets for developing drugs for treating
PT   Alzheimer's disease, ischemia or cerebral infarction, or for diagnosing
PT   these diseases -
XX
PS   Example 1; Fig 1; 53pp; English.
XX
CC   The present invention describes a neural thread protein (NTP) peptide
CC   having an amino acid sequence selected from ABB81511 to ABB81529 and
CC   their homologues, which are referred collectively as Harl1l peptides (I).
CC   (I) have neuroprotective, nootropic, vasotropic and cerebroprotective
CC   activities, and can be used in peptide therapy. The Harl1l peptide
CC   sequences can be used as analogues for NTP in therapeutic or diagnostic
CC   peptides by replacing NTP with the peptide in such an assay. The Harl1l
CC   peptides are also useful as a trap material in a diagnostic or
CC   therapeutic assay. Therefore, the Harl1l peptides are useful in
CC   assays, protein and antibody purification, therapeutics or diagnostics.
CC   In particular, the peptides are also useful for diagnosing Alzheimer's
CC   disease, Down's syndrome, neuroectodermal tumours, astrocytomas,
CC   glioblastomas, hypoxia, ischaemia or cerebral infarction. The peptides
CC   are also useful as targets for drug development for the treatment of
CC   these diseases. The present sequence encodes a neural thread protein
CC   given in the exemplification of the present invention.
XX
SQ   Sequence 1442 BP; 314 A; 402 C; 311 G; 415 T; 0 other;

Query Match             100.0%;  Score 1442;  DB 24;  Length 1442;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1442;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy      1 TTTTTTTTTTTGAGATGGAGTTTTTCGCTCTCTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTGAGATGGAGTTTTTCGCTCTCTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60

Qy     61 TCTCGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGAATTCTCCTGCCTCAGCCTCCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGAATTCTCCTGCCTCAGCCTCCC 120

Qy    121 CAGTAGCTGGGATTACAGGCATGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTTTAG 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCATGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTTTAG 180

Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240

Qy    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCGCGCCCGGCC 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCGCGCCCGGCC 300

Qy    301 TGGCTAATTTTTGTGTGTTTAGTAGAAACAGGGTTTCACTGTGTTGGTCTCCTGAGC 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCTAATTTTTGTGTGTTTAGTAGAAACAGGGTTTCACTGTGTTGGTCTCCTGAGC 360

Qy    361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGC 420

Qy    421 CTGGCCTTTTTATTTTTATTTTTTAAGACACAGGGTGTCCCACTCTTACCCAGGATGAAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGCCTTTTTATTTTTATTTTTTAAGACACAGGGTGTCCCACTCTTACCCAGGATGAAG 480

Qy    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTGAGATCAAGCAATCTCCTG 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTGAGATCAAGCAATCTCCTG 540

Qy    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAATTTTA 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAGACATGCACCACTACACCTGGCTAATTTTA 600

Qy    601 TTTTTATTTTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTATTTTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT 660

Qy    661 GGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCCTGCCCA 720

Qy    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTGTATTTT 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTGTATTTT 780

Qy    781 TAGTAGAGATGGGGTTTCACCGTTGTTGGCCAGGTTGGTCTCGAACTCCTGACCTCAGGT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGATGGGGTTTCACCGTTGTTGGCCAGGTTGGTCTCGAACTCCTGACCTCAGGT 840

Qy    841 TGCCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCTTA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCTTA 900

Qy    901 TTTTTAATTTTTGTTTGTTTGTTTTGAAATGGAAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTTAATTTTTGTTTGTTTGTTTTGAAATGGAAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960

Qy    961 GGCCAAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCAAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020

Qy   1021 GCCTCCCAAGCAGCTGGGATTACGGGCACACCCGGCCGCTAATTTTTGTATTTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCAGCTGGGATTACGGGCACACCCGGCCGCTAATTTTTGTATTTT 1080

Qy   1081 TCATTAGAGGCGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGGCGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140

Qy   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCAG 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCAG 1200

Qy   1201 CCGGCTAATTTAGATAAAAAAAATATGTAGCAAATGGGGGTCTTGCTATGGTTGCCCAGGCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGCTAATTTAGATAAAAAAAATATGTAGCAAATGGGGGTCTTGCTATGGTTGCCCAGGCT 1260

Qy   1261 GGTCTCAAACTTCTGGCCTTCATGCAATCCTTCCAAATGAGCCAACAACACCCAGCAGTCA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCAAACTTCTGGCCTTCATGCAATCCTTCCAAATGAGCCAACAACACCCAGCAGTCA 1320

Qy   1321 CATTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATACAATAAAACATGT 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATACAATAAAACATGT 1380

Qy   1381 CAAACCTGCAAAATTCAGTAGTAACAGAGAGTTCTTTATTTAAACAAAAGCTTTAGAG 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAAATTCAGTAGTAACAGAGAGTTCTTTATTTAAACAAAAGCTTTAGAG 1440

Qy   1441 CA 1442
          ||
Db   1441 CA 1442
```

```
RESULT 2
AAT27738
ID   AAT27738 standard; DNA; 1418 BP.
XX
AC   AAT27738;
XX
DT   13-NOV-1996  (first entry)
XX
DE   Neural thread protein coding sequence.
XX
KW   Neural thread protein; NTP; diagnosis; detection;
```

KW Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KW monoclonal antibody; binding fragment; ds.
XX
OS Homo sapiens.
XX
FH
FT Key   Location/Qualifiers
FT CDS   14..1207
FT   /*tag= a
FT   /product= Neural thread protein.
XX
PN (GEHO) GEN HOSPITAL CORP.
XX
PA De LA MONTE S, Wands JR;
XX
PI WO9615272-A1.
XX
PD 23-MAY-1996.
XX
PF 14-NOV-1995; 95WO-US17111.
PR 14-NOV-1994; 94US-0340426.
XX
DR WPI; 1996-259865/26.
DR P-PSDB; AAR95913.
XX
PT Detection of neural thread protein in diagnosis of Alzheimer's
PT disease - also NTP DNA and protein sequences used in gene and
PT anti:sense therapy
XX
PS Claim 24; Page 168-170; 238pp; English.
XX
CC A method for detecting the presence of neural thread protein (NTP)
CC having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC subject comprises (a) contacting a sample from a human subject that
CC is suspected of containing the NTP with at least one molecule
CC capable of binding to the protein; and (b) detecting any of the
CC molecule bound to the protein. The binding molecule is selected
CC from an antibody free of natural impurities, a monoclonal antibody
CC or a binding fragment of either of these. The method may be used for
CC diagnosing the presence of Alzheimer's disease, neuroectodermal
CC tumours and a malignant astrocytoma in a human.
XX
SQ Sequence 1418 BP; 302 A; 396 C; 315 G; 405 T; 0 other;

Query Match   84.8%; Score 1223.4; DB 17; Length 1418;
Best Local Similarity   96.9%;
Matches 1375; Conservative   0; Mismatches   31; Indels   13; Gaps   12;

Qy   2 TTTTTTTTTTTTGAGATGGAGAGAGTTTTCGCTCTTGTTGTCCCAGGCTGGAGTGCAGCGAAT 61
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTCCCAGGCTGGAGTGCAATGGCGCAAT 60

Qy   62 CTCAGCCTCACCGGAACCTCCGCCTCCGCCTCCCAAAGCGATTCCTCGCCTCAGCCTCCC 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTCAGCCTCACCGAACCTCCGCCTCCGGGTTCAAGCGATTCTCCTCAGCCTCAGCCTCCC 120

Qy   122 AGTA-GCTGGGATTACAGGCGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTTTTTAG 180
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AGTAGGCTGGGATTACAGGCATGGTGCA-CCACGCTCGGCTAATTTTTGTATTTTTTTTTAG 179

Qy   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   180 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 238

Qy   241 CCTCCGTCTCCGGCCTCCCAAAGTGCTGAGATTACAAGGACTGGGCAACCACCGCCTGCC 299
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   239 CTCCCGTCTCCGGCCTCCCAAAGTGCTAGATTACAGGACTGGGCAACCACCATGCCCGGC 298

Qy   300 CTGGCTAATTTTTGTGTTAGAAAACAGGGTTTCACTGAATACATGATTGTCTCCTGAG 359
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   299 CTGGCTAATTTTTGTGTTAGAGAAACAGGGTTTCACTGATGTTGGTCAGGCTGGTCTCCTGAG 358

Qy   360 CTCAAGCAGTCCACCTCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTG 419
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   359 CTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGGGTGCAGCCGTG 418

Qy   420 CCTGGCCTTTTTTATTTTTTATTTTTTTAAGCACTACAGGTGTCCCACTCTTACCCAGGATGAA 479
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   419 CCTGGCCTTTTTATTTTTATTTTTTTAAGACACACAGGTGTCCCACTCTTACCCAGGATGAA 478

Qy   480 GTGCGAGTGGTGTGGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCATCCTCCT 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 GTGCAGTGGTGTGGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAGATGCA 537

Qy   540 GCCTCAGCCTCCC -AAGTAGCTGGGACCAAAGACAATGCACCACCACTACACCTGGCTAATTTT 598
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   538 GCCTCAGCCTCCCAAAGTAGCTGGGACCAAAGACAATGCACCACCACTACACCTGGCTAATTTT 597

Qy   599 GTGCGAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCC 658
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   598 TATTTTTATTTTTTTTAATTTTTGAGACAGAGTCTCACCCAGGCTGGAGTGCA 657

Qy   659 GTGGCGCAAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCCTGCC 718
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   658 GTGGCGCAAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCCTGCC 717

Qy   719 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCCGGCTAATTTTTTGTATTT 778
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   718 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCCAGCTAATTTTTTGTATTT 777

Qy   779 TTTAGTAGAGATGGGG-TTCACCAGGTTGGTCGGCCAGGTTGAT-CTTGATCTCTGGACCTTGT 836
          ||||||||||||||| ||||||||||||||||||||||||||| |||||||||||||||||
Db   778 TTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCTTGACCTTGT 837

Qy   837 GATCTGCCTCGCCTCGCGCCTCCAAAGTGCTGGGATTACAGG-CGTGAGCCACCACGCCCG 895
          |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   838 GATCTCGCCTGCCTCGGCCTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCG 897

Qy   896 GCTTAATTTTTTAATTTTGTTTGTTTGTTTGAAATGGAATCACTTGTTTACACTCGTTGGAGT 955
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   898 GCCTATTTTTTAATTTTGTTTGTTTGTTTGAAATGGAATCACTTGTTTACACTCGTTGGAGT 957

Qy   956 GCAATGGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGGCTCAAGCGGATTCTCCTG 1015
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   958 GCAATGGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGCATTCTCCTG 1017

Qy   1016 TCTCAGCCTCCCAAGCAGCTGGGATTACGGGGCGACCACCACCCCGGCTAATTTTG 1075
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1018 TCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCACCACCCCGGCTAATTTTG 1076

Qy   1076 TATTTTCAGTTAGAGGCGGGGTTTCACCAGGTTGTCAGGCTGGTCTCAAAACTCCTGACCTT 1135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1077 TATTTCAGTTAGAGGCGGGGTTTCACCAGGTTTCAGGCTGGTCTCAAACTCCTGACCT 1136

Qy   1136 CAGGTGACCCAGCCACCGCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTCA 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1137 CAGGTGACCCAGCCGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCGTGA--CGCCTCA 1194

Qy   1196 CCCAGCCGCGGCTAATTTAGTAATAAAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCCC 1255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1195 CCCAGCCGGGGCTAATTTTAGTAATAAAAATATGTAGCAATGGGGGGTCTTGCTATGTTGCC 1254

Qy   1256 AGGCTGGTCTCAAACTTCTCGGCTTCATGCAAATCCTTCCAAATGAGCCACCACACCCGGCC 1315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1255 AGGCTGGTCTCAAACTTCTGGCTTCATGCAATATCCTTCCAAATGACCCACACACCCAGCC 1314

Qy   1316 AGTCACATTTTTTAAACAGTTTACATCTTTATTTTAGTATATACTAGAAAAGTAATACAAATAA 1375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1315 AGTCACA-TTTTTAAACAGTTTACATCTTTATTTTAGTATATACTAGAAAAGTGATACAGGATAAC 1373

Qy   1376 CATGTCAAAACCTGCAAAATTCAGTAGTAACACAGAGTTCTTT 1414
          ||||||||||||||||||||||||||| |||||
Db   1374 ATGGGCGGAAACCTGCAAAATTCGAGTAGTAGTACAGAGTCTTT 1412

RESULT 3

```
AAQ77883
ID   AAQ77883  standard; cDNA; 1381 BP.
XX
AC   AAQ77883;
XX
DT   06-JUL-1995  (first entry)
XX
DE   Neural thread protein AD10-7 cDNA.
XX
KW   Neural thread protein AD10-7; Alzheimer's; neuroectodermal tumours;
KW   malignant astrocytomas; glioblastomas; ss.
XX
OS   Rattus rattus.
XX
PN   WO9423756-A.
XX
PD   27-OCT-1994.
XX
PF   20-APR-1994;  94WO-US04321.
XX
PR   20-APR-1993;  93US-0050559.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PI   De LA MONTE SM, Wands JR;
XX
DR   WPI; 1994-341497/42.
XX
PT   Detection of neural thread proteins - to detect sporadic and
PT   familial Alzheimer's disease, neuroectodermal tumours, malignant
PT   astrocytomas and glioblastomas (Eng).
XX
PS   Example 4; Fig 16R; 158pp; English.
XX
CC   AAQ77883 is the AD10-7 neural thread protein (NTP) cDNA. This
CC   sequence was used in the development of an antibody dependent
CC   method, for the detection of NTPs. This new method could be
CC   used to diagnose Alzheimer's disease (differentiating between
CC   sporadic and familial), neuroectodermal tumours, malignant
CC   astrocytomas and glioblastomas.
XX
SQ   Sequence 1381 BP; 296 A; 385 C; 301 G; 399 T; 0 other;

Query Match             74.8%;  Score 1078.6;  DB 15;  Length 1381;
Best Local Similarity   94.3%;
Matches  1316;  Conservative    0;  Mismatches   54;  Indels   25;  Gaps   18;
                                    Pred. No. 1.4e-277;
```

```
Qy      2  TTTTTTTTTTTTGAGATGGAGGAGTTTTCGCTCTCTTGTTGCCCAGTGCGGAGTGCAATGGCGCAAT     61
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  TTTTTTTTTTTGAGATGGAGTTTTCGCTCTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT          60

Qy     62  CTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCC          121
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCC          120

Qy    122  AGTAGCTGGGATTACAGCATGTGCACCCACCGCTCGGCTAATTTTGTAATTTTTAGT            181
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  AGTAGCTGGGATTACAGGCATGTGCA-CCACGCTCGGCTAATTTTGTATTTTTTAGT            179

Qy    182  AGAGATGGAGTTT--CTCCAATGTTGGTCAGCGTGGTCTCGAACTCCCGACCTCAGATGAT        239
           ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db    180  AGAGATGGAGTTTAACTCCAATGTTGGTCAGCTGGTCTCGAACTCCCGACCTCAGATGAT        239

Qy    240  CCCTCCGTCTCGGCCTCCCAAAGTGCT---AGATACAGGACTGCCACCCATGCCCGGCCT        295
           ||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db    240  CTCCAGCTCACCACGCAACCTCCGGTCCCAAAGTGCTGAGATTACAGGCATGAGCCACCA        299

Qy    296  AGTAGCTGGGATTACAGCATGTGCACCCACCGCTCGGCTAATTTTGTCACTGATG-TGCC        354
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    300  CTGCCTGGCCTAATTTTTGTGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTC        359

Qy    355  CTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAG        414
           |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    360  CTCAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG     418

Qy    415  CCGTGCCTGGCCTTTTTATTTTTTATTTTTTTGAGACAGAGGTGTCCCACTCTTTACCCAGG        474
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    419  CCGTGCCTGGCCTTTTTATTTTTTATTTTTTTGAGACACAGGTGTACCACTCTTTACCCAGG     478

Qy    475  ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCACAAGC-AT        533
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    479  ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTGAACTCCTGAGGATCAAGCAAT     538

Qy    534  CCTCCTGCCTCAGCCTCCCAAAGTAGCTGGGACCAAAAGACATGCAACTACTACCTGGGCTA        593
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539  CCTCCTGCCTCAGCCTCCCAAAGTAGCTGGGACCACAACATCACCACTACCACCTGG-TA     597

Qy    594  ATTTTTATTTTTTATTTTTTTGAGACAGAGGGTTTCACCATGTTACCCAGGCTGGA        653
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598  ATTTTTATTTTTTATTTTTTTTGAGACAGAGGTGTACCACTCTTTACCCAGG     656

Qy    654  GTGCAGTGGCGCAATCTCTGCCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC        713
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    657  GTGCAGTGGCGCGAATCTCTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC     716

Qy    714  TGCCCCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCACCACCGCGCTAGCTAATTTTTTT        773
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    717  TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCGCGCTAGCTAATTTTTTT     776

Qy    774  GTATTTTTAGTAGAGATGGGG--TTCACCATGTTGGCCAGGTTGATCTTGATCTCTGGACC        832
           ||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db    777  GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATCTTGATCTCTGGACC     836

Qy    833  TTGTGATCTGCCTCGGCCT-CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCACG        891
           |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    837  TTGTGATCTGCCTGCCTCACCCAAAGTGCTGGGATTACAGG-GTGAGCCACCAC          894

Qy    892  CCCGGCCTTATTTTTTTGTTTTGTTTTGTTTTTGAGATGGAATCTCACTCTGTTACCCAGGCTG        951
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    895  GCCGGCCTAATTTTTAATTTTTGTTTTGTTTTTGAGAATCTCACTCACTCTGTTACCCAGGCTG     954

Qy    952  GAGTGCAAGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGGTCAAGCGATTCT        1011
           |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    955  GAGTGCAATGG-CAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGG-TCAAGCGATTCT     1012

Qy   1012  CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCCACCACCACCCCGGCTAATT        1070
           ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   1013  CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGG----ACCTGCACCACCACCCCGGCTAATT     1071

Qy   1072  TTTGTATTTTCATTAGAGACGGGGTTTCACCATGTTGTCAGGCT-GGTCTCGAACTCCT        1130
           ||||||||||||||||||||||||| |||||||||||||||||
Db   1071  TTTGTATTTTCATTAGAGGGGG--TTTACCATATTGTGGGCTTGGTCTCAAACTCCT          1130

Qy   1131  GACCTCAGGTGACCCAGCCGCCTCAGCGTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA        1190
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129  GACCTCAGGTGACCCACCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA     1188

Qy   1191  CCTCACCCAGCCGCGGCTAATTTGATAAAAAAATATGTAGAGAATGGGGGTGTCTTGCTATGT        1250
           ||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db   1189  CCTCACCCAGCCGCGGCTAATTTTGGAATAAAAAAAATATGTAGAATGGGGGTCTTGCTATGT     1246

Qy   1251  TGCCCCAGGCTGGTCTCGAACTCCTCAGGCTCAAGCAATCTTCCAAATGAGCCACCACACC        1310
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1247  TGCCCCAGGCTGGTCTCAGGGAATTAGTCCATTCTCCTTCAGCTTCCAAAATGAGCCACACACC     1310

Qy   1311  CAGCCAGTCACATTTTTTTAAAACAGTTACACTTTATTTAGTATATACTAGAAAAGTAAATACA        1370
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1307  CAGCCAGTCACATTTTTTTAAACAGTTACAACTTTATTTTAGTATACTAGAAAAGTAATACA     1366

Qy   1371  ATAAACATGTCAAAC     1385
           |||||||||||||
Db   1367  ATAAACATGTCAAAC     1381
```

```
RESULT 4
AAT27765
```

```
ID   AAT27765 standard; cDNA; 1381 BP.
XX
AC   AAT27765;
XX
DT   14-NOV-1996  (first entry)
XX
DE   AD 10-7 human neural thread protein clone (partial sequence).
XX
KW   Neural thread protein; NTP; diagnosis; detection;
KW   Alzheimer's disease; neuroectodermal tumour; malignant astrocytoma;
KW   monoclonal antibody; binding fragment; ds.
XX
OS   Homo sapiens.
XX
PN   WO9615272-A1.
XX
PD   23-MAY-1996.
XX
PF   14-NOV-1995;   95WO-US17111.
XX
PR   14-NOV-1994;   94US-0340426.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   De LA MONTE S, Wands JR;
XX
DR   WPI; 1996-259865/26.
XX
PT   Detection of neural thread protein in diagnosis of Alzheimer's
PT   disease - also NTP DNA and protein sequences used in gene and
PT   anti:sense therapy
XX
PS   Example 4c; Figure 16r; 238pp; English.
XX
CC   A method for detecting the presence of neural thread protein (NTP)
CC   having a molecular weight of 8, 14, 17, 21, 26 or 42 kD in a human
CC   subject comprises (a) contacting a sample from a human subject that
CC   is suspected of containing the NTP with at least one molecule
CC   capable of binding to the protein; and (b) detecting any of the
CC   molecule bound to the protein. The binding molecule is selected
CC   from an antibody free of natural impurities, a monoclonal antibody
CC   or a binding fragment of either of these. The method may be used for
CC   diagnosing the presence of Alzheimer's disease, neuroectodermal
CC   tumours and a malignant astrocytoma in a human. A number of clones
CC   of neural thread protein were isolated from healthy 17-18 week old
CC   foetal human brain (HB) 2 year old temporal lobe neocortex and end
CC   stage Alzheimer's disease (AD) cerebral cortex. See AAT2753-75.
XX
SQ   Sequence 1381 BP; 295 A; 385 C; 301 G; 400 T; 0 other;

Query Match       74.8%;  Score 1077;  DB 17;  Length 1381;
Best Local Similarity  94.38%;
Matches 1315;  Conservative    0;  Mismatches  55;  Indels  25;  Gaps  18;
```

```
Qy      2 TTTTTTTTTTTTGAGATGGAGATTGGAGAGTTTTTCGCTCTTGTTGTTGCCGACCTTTGGTGTTGCCGACCGGACCGGCCGCAAT 61
          ||||||||||| |||||||||| |||||||||||||||||||||| ||||            ||||||||||||||||||
Db      1 TTTTTTTTTTTGAGATGGAGAGTTTTTCGCTCTTGTTGCCCAGGTGCAATGGCGCCAAT 60

Qy     62 CTCAGCTCACCGACAACCTCCGCGCTCCGGGGTTCAACCGATTCTCCAGCCTCCCC 121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGACAACCTCCGCGCTCCGGGGTTCAACCGATTCTCCAGCCTCCCC 120

Qy    122 AGTAGCTGGGATTACAGGCANGTGCACCACCACGCTCGGCTAATTTTGTATTTTTTAGT 181
          ||||||||||||||||||||||||||  ||||| ||||||||||||||||||||||||
Db    121 AGTAGCTGGGATTACAGGCANGTGCA--CCACCTCGGCTAATTTTGTATTTTTTAGT 179

Qy    182 AGAGATGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT 239
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 AGAGATGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGAT 239

Qy    240 CCCTCCGCGGCTCGGCCGCCTCCCCAAAGTGCT---AGATACAGGACTGGCCCACCATGCCCGG-CT 295
          | ||||||||||||||||||| || |||||||||    ||| ||| | ||||||||||||||||||| ||
Db    240 CTCCGCGGCTCGGCCGCCTGCCCAAAGTGCTCTGAGATTACACGGCCATGAGCGCCACCATGCCCGGGCT 299
```

```
Qy    296 CTGGCCTGGCTAATTTTTGTGGTAGAAANACAGGGTTTCACTGATG-TGCCCAAGCTGGTCTC 354
          ||| |||||||||||||||||||||||| ||||||||||||||||| ||||||||||||||||
Db    300 CTGCCTGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTTGCCCAAGCTGGTCTC 359

Qy    355 CTGAGCTCAAGCAGTCCACCGCCTCCCAGCCTCCCAAGTGCTGGGATTACAGGCGT-CAG 414
          |||| ||||||||||||||| || || ||||||||||||||||||||||||||||| |||
Db    360 CTGACCCTCAAGCAGTCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAG 418

Qy    415 CCGTGCCTGGCCGTTTATTTATTTATTTATTTTTTTGAGACAGAGTGTCCCACTCTTACCCAGG 474
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    419 CCGTGCCTGGCCGTTTATTTATTTATTTATTTTTTTGAGACAGAGTGTCCCACTCTTACCCAGG 478

Qy    475 ATGAAGTGCCAGTGGTGTGATCAACAGCTCACTGCAGCCTCAACTCCTGAGATCAAGCAAT-AT 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    479 ATGAAGTGCCAGTGGTGTGATCAACAGCTCACTGCAGCCTCAACTCCTGAGATCAAGCAAT 538

Qy    534 GTGGCAGTGGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC 593
          
Db    539 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACCACGTACCTGGCTA 597

Qy    594 ATTTTTATTTTTATTTTTATTTTTTGAGACAGAGTCGAAACTGTCACCCAGGCTGGA 653
          |||||| |||||||||||||||||||||||||||||  |||||||||||||||||||
Db    598 ATTTTTTATTTTTATTTTTATTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGGA 656

Qy    654 GTGCAGTGGCGCAAATCTTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC 713
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    657 GTGCAGTGGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC 716

Qy    714 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTT 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    717 TGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCAGCTAATTTTTTT 776

Qy    774 GTATTTTTAGTAGAGATGGGGG-TTCACCATGTTCGCCAGGTTGATCTCGATCTCTTGACC 832
          ||||||||||||||||||||| ||||||||||| || ||||| |||||||||||||||||||
Db    777 GTATTTTTAGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGATCTCTTGACC 836

Qy    833 TTGTGATCTGCCTCGCCTC-CCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCACG 891
          |||||||||| ||||||||  ||||||||||||||||||||||  | | || |||
Db    837 TTGTGATCTGGCCTCGCCTGCCCAAAGTGCTGGGATTACAG--GTCGTGACTCCAC 894

Qy    892 CCCGGCGTTATTTTTATTTTTGTTGTTGTTGTTGAAATGGAATCTCACTCTGTTACCCAGGCTG 951
          ||||| | |||||| |||||||| ||| ||| |||||||||||||||||||||||||||||| | 
Db    895 GCCGGCCTATTTTTTAATTTTTGTTTGTTTGTTGAAATGGAATCTCACTCTGTTACCCAGGTCG 954

Qy    952 GAGTGCAATGGCGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCT 1011
          |||||||||||| ||||||||||||| | ||||||| |||||||||||| ||||||| |||||
Db    955 GAGTGCAATGG-CAAAYCTCGGCTACYCGCAACCTTCGCCTCCCGGG-TCAAGCGAATTCT 1012

Qy   1012 CCTGTCTCAGCCTCCCAAGCACGTCGGGATTAGGGGCCACCTGCCACCACCACCCGCTAATT 1071
          |||||||||||||||||||||| || |||||| ||||  |||||||||||| || ||||||||
Db   1013 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGG--ACCTGCCACCACACCCCGCTAATT 1070

Qy   1072 TTTGTATTTTCATTAGAGGCGGGGGTTCACCATGTTGGTCAGGCTGGTCTCAAACTCCT 1128
          ||||||||||||||||||||||    ||| ||||||||||||||||||||||||||||||
Db   1071 TTTGTATTTTCATTAGAGGCGGG---TTTACCATGTTGGTCAGGCTGGTCTCAAACTCCT 1128

Qy   1131 GACCTCAGGTGACCGACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1190
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 GACCTCAGGTGACCGACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1188

Qy   1191 CCTCACCCAGCCNGGCTAATTTGGAATAAAAAAAATATGTAGGCAATGGGGGGTCTTGTATGT 1250
          |||||||||||| ||||||||||||||   ||||||||||||||||||||||  |||||||||||
Db   1189 CCTCACCCAGCCGGCTAATTTGGAAT---AAAATATGTAGGCAATGGGGG--TCTTGTATGT 1246

Qy   1251 TGCCCAGGCTGGTCTCAAACTTCTCGGCTTCATGCAATCCTTCCAAATGAGCCAACACAC 1310
          |||||||||||||||||||||||| |||||||  ||||||||| | |||||||||| ||||
Db   1247 TGCCCAGGCTGGTCTCAAACTTCTGGCTTCAGTCAATCCTTCAGAAATGAGCCACAACAC 1306

Qy   1311 CAGCCAGTCACATTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATACA 1370
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   1307 CAGCCAGTCACATTTTTTAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATACA 1366
```

Qy   1371 ATAAACATGTCAAAC 1385
           ||||||||||||||||
Db   1367 ATAAACATGTCAAAC 1381

RESULT 5
ID   ABL62910/c
ID   ABL62910 standard; DNA; 65608 BP.
XX
AC   ABL62910;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Breast cancer related gene sequence SEQ ID NO:1247.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209551P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   26-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   27-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   28-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236034P.
PR   28-SEP-2000; 2000US-236109P.
PR   29-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-268891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   03-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.

PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
XX
PS   Claim 1; SEQ ID 1247; 44pp; English.
XX
CC   The present invention describes a method (MI) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. M1 has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 65608 BP; 15987 A; 16979 C; 16761 G; 15878 T; 3 other;

Query Match   40.8%;   Score 587.8;   DB 24;   Length 65608;
Best Local Similarity 70.6%;   Pred No.5.4e-146;
  Matches 940; Conservative   0; Mismatches 347; Indels   44; Gaps 10;

Qy   4 TTTTTTTTGAGATGGAGTTTTTCGCTCTTGTTTGCCCAGGCTGGAGTGCAAATGGGCAATCT 63
          |||| |||||||||||| ||||||||||||||||||||||| |||| |||| ||
Db 39993 TTTTGTTTTGAGATGGAGTTTT-GCTCCTTGTTGCCCAGGCTGAAGTGTAATGGCACGATCT 39935

Qy   64 CAGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGAATTCTCCGCCTCAGCCTCCCAG 123
          ||||||||||||| |||||||||||||| |||||||||||||||||||||| | |||
Db 39934 CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAGCAATTCTCCTACCTCAGCCTCCCGAG 39875

Qy   124 TAGCTGGGATTACAGGCATGTGCACCCACGCTCGGGCTAATTTTGTATTTTTTTAGTAG 183
          |||||||||||||||||||||| ||||| || ||||||| ||||||||||| ||||
Db 39874 TAGCTGGGATTACAGGCATGTGCCACCACCGCCTAATTTTTGTATTTTT---AGT 39819

Qy   184 AGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCT 243
          || | ||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db 39818 AGACGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACCCCCGACCTCAGATGATCCGC 39759

Qy   244 CCGTCTTCGGGCCTCCCAAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCCTGCCTGG 303
          ||||||||||| ||||||||||| ||||||||| || |||||| |||| ||||| ||||
Db 39758 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGTGTGAGTTTTACAAGAAAACATTTA 39700

Qy   304 CTAATTTTTTGTGTAGAAACAGGGTTTTCACTGATGTGCCCAAGCTGGTCTCCGAGCTCA 363
          | || ||||||| |||||| ||||| ||||||||||||||||||||||||||||||
Db 39699 AGTATGAAAAGCAGTACCTCCGGAATATGCACGGCGCCTCCAACCCCAGCCTCTGAAGACTC 39640

Qy   364 AGCAGTCACCTTGCCTCAGCCTCCCAAAAGTGCTGGGATTA------------ 403
          ||||||||||||||||||||| ||||| ||||||||||||
Db 39639 TTCACTTGTTCCCTGGCCTGGGCCACCACCTCACCTAAAACTGGGCGGTGTTCTCTT 39580

Qy   404 CAGGCGTTGCAGCCGTGCCTGGCCTTTTTATTTATTTTATTTTTTGAAGACACAGGTCCCAC 463
          ||||||| |||||| ||||||| ||||| |||||||||||||||| ||
Db 39579 CCCCCCTCCATCCTGCCTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCTTGTT 39520

Qy   464 TCTTACCCAGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCAACTCCTGA 523
          ||| || | ||| ||| ||||| |||||| ||||||||||||||||||||||||||||
Db 39519 TGTTGCCCAGGCTGGAGTGCAGTGGCGCAAAACATGGCTCACTGCAGCCTTGACCTCCTGG 39460

Qy   524 GATC-AAGCATCCTCTGCCCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACACT 582
          || | ||||||| |||||||||||||||||||||||||||||| ||| || |||||||
Db 39519 TGTTGCCCAGGCTGGAGTGCAGTGGCGCAAAACATGGCTCACTGCAGCCTTGACCTCCTGG 39460

```
Db  39459  GCTCAAAGGATCCTCCCACCTTCCACCTTCCAGCCTCCTGAGTGAGTAGCTGGGACTATAGGTGCATGCCACC  39400
Qy    583  ACACCTGGCTAATTTTTATTTTTATTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCA  642
Db  39399  ATGCCCAACTAATTTTTTTTTTT-------TTTTTTGAGATGGAGTTTCGCTCTGTTG  39348
Qy    643  CCCAGGCTGGAGTGCAGTGGCGCCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCA  702
Db  39347  CCCAGGCTGGGGTGGTGCAAATGGTGCGATCCTGCTCACTGCAACCTCCGTCCTCGGGTTCA  39288
Qy    703  AGTTATTCTCTGCCCCAGCCTCCTGAGTAGCTGGGACTACAAGGCGCCCACCACGCCTAG  762
Db  39287  AGCAATTCTCCTGCCTCAAACCCCCAAGTAGCTGGGATTACAGGCATGCCATCACACC  39228
Qy    763  CTAATTTTTTTGTATTTTTAGTAGAGATGGGG-TTCACCATGTTGGCCAGTTGATCTTG  821
Db  39227  TGGCTAATTTTTGTATTGTTAGTAGAGAATGGGGTTTCACTCAGGGTGTGTTCTCA  39168
Qy    822  ATCTCTGGACCTT--TGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG  879
Db  39167  AACTCCTGGACCTCAGCGAGAGATCCTCCGCCTCGGCCTCCCAAGGCGTGGGATTACAGGTG  39108
Qy    880  TGAGCCACCACCCCCGGCCTT--ATTTTTAATTTTTTGTTTGTTTGTTTGAAATGGAATCTCACTC  937
Db  39107  TGAGCCACTGTGGCTGGCTGGCCTGGCCCAATTTTTTTTTTTTTTTTTTTTTTGGAGACAAAGTCTCGC  39048
Qy    938  TGTTACCCAGGCTGGAGTGCAAATGGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGG  997
Db  39047  TCTTTCCCAGGCTGGGAGTGGATGGCACCACCGTCGCGGCTCACTGCGAACCTCTGCCTTTCGG  38988
Qy    998  GCTCAAGCGGATTCTCCTGTTCTCAGCCTCCCCAAGCAGCGTGGGATTTACGGGCCACCTGCCACC  1057
Db  38987  GTTCAAGCGGATTCTCCTGCCTCAGTCTCCCAAGTAGCTAGGACTACAGGCGCATGGTGCCCACC  38928
Qy   1058  ACACCCCGCTAATTTTTGTATTTTTTCATTTAGAGAGGGGGGGTTTCACCCATATTTGTCAGGCTG  1117
Db  38927  ACGCCCGGCTAATTTTTGTATTTTTTAGTAGAGACGAGGTTTTCACCATGTTGGCCAGACTG  38868
Qy   1118  GTCTCAAACTCCTGACCTCAGGTGACCCACCTGCCTCAGCCTCCAAAGTGCTGGGATTA  1177
Db  38867  GTCTTGAACTCCTAACTTCAGGTGATCCACCCGCCGTCGGCCTCCCAAAGTGCTGGGATTA  38808
Qy   1178  CAGGCGTGAGCCACCGCACCACCTCACCCAGCC----GGCTAATTTAGATAAAAAATATGTTGAGCAAT  1233
Db  38807  CAGGCATGAGCCACCACCACCTGGCCTGTCGCCACTAATTAAAAAAAAATTTTGTAGAG  38748
Qy   1234  GGGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAACTTCTGGGCTTCATCCAATCCTTCC  1293
Db  38747  ACAGGACGATTCCACTATTGTTCTCAGACTGGTCTTGAACTCCTGGGCTTAAGGGATCCTTCT  38688
Qy   1294  AAATGAGCCAC  1304
Db  38687  GTCTCAGCCTC  38677
```

```
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 65608 BP; 15987 A; 16979 C; 16761 G; 15878 T; 3 other;

Query Match    40.8%;  Score 587.8; DB 24; Length 65608;
Best Local Similarity   70.6%;  Pred. No. 5.4e-146;
  Matches  940; Conservative   0; Mismatches 347; Indels  44; Gaps  10;

Qy       4 TTTTTTTTTTGAGATGGAGTTGTTTCGTCTGTTGTTGCCCAGGCTGGAGTGGCGCAATCT 63
           ||||| |||||||| ||||||||||||| | ||||||||||||||||| ||| |  ||||
Db   39993 TTTTTGTTTTGAGAATGGAGAGTTTTT-GCTCTTGTTGCCCAGGCTGAAGTGTTAATGGCACAGAATCT 39935

Qy      64 CAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG 123
           |||||||||  |||||||||||| |||||||| ||||| |||||||||||||||||||| ||
Db   39934 CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAGCAAATTCTCCTACCTCAGCCTCCCGAG 39875

Qy     124 TAGCTGGGATTACAGGCACAGTGCCACCACGCCTCGGCTAATTTTTGTATTTTTTTAGTAG 183
           |||||||||||||||||| | |||||||||| ||  ||||||||||| |||||| | |||||
Db   39874 TAGCTGGGATTACAGGCACATGTGTCACCACGCCTCGGCTAATTTTTTGTATTTTT----AGT 39819

Qy     184 AGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCCT 243
           || |||||||||||||||| |||||| || |||| ||||||  |||| || |||| ||||
Db   39818 AGACGGGGTTTCTCCATGTTGGTCAGGCTGGTCTTGAACCCCCGACCTCAGATGAATCCGC 39759

Qy     244 CCGTCTCGGCCGCCTCCCCAAAGTGCTGAGATTACAGGCGTGGCACCACGCCCGGCC 303
           | || ||||||||||| |||||| ||||||||||||||  |||||||| ||| |||||
Db   39758 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTACAAGAAAACATTTA 39700

Qy     304 CTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCTGAGCTCA 363
           ||||||||||| ||  | ||||||||||||||| ||||||| || |||| |||||||| |
Db   39699 AGTATGAAAAGCAGTAACTCCGGAATATGCACGCCCGTCCCCAGCCTGTCCTGAAGACTC 39640

Qy     364 AGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTA 403
           |||||||||||| |||||||||||||||||  |||||||| |
Db   39639 TTCACTTGTCCTGGGCCCACCACCTCCACCTAAAACTGGGCTGTTGTTCTTT 39580

Qy     404 CAGGGCGTGCACCGGTGCCTCGGCCTTGTTTTATTTATTTAAGACACAGGGTGTCCCAC 463
           ||||| ||||||||| ||||||| |||||||||||||||| |||||||||| ||||| |
Db   39579 CCCCCTCATCCCTGCCTTTTTGTTTTTTTTTTTTTTTTTTGAGACAGAGTCTTGTTT 39520

Qy     464 TCTTTACCCAGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCACCTTCAACTCCTGA 523
           || |||||||||| ||||||| ||||| | ||||| ||||| ||||| |||| |||||||
Db   39519 TGTTGCCCAGGCTGGAGTGCAGTGGCGCAAACATGGCTCACTGCAGCCTCGAATCCTGGG 39460

Qy     524 GATC-AAGCAATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACACATGCACCACT 582
           |||| ||||||||||||||||||||||||||| | ||||||||||||||| ||||||| |||
Db   39459 GCTCAAAAGGATCCTCCCACCTTCAGCCTCCCGAGTAGCTGGGACTATAGGTGCATGCCACC 39400

Qy     583 ACACCTGGCTAATTTTTTATTTTTTAGTTTTAATTTTTTTTGAGACAGAGTCTCAACTCTGTCA 642
           |||||| ||||||||||| ||||| |||||||||||||| ||| |||||| ||| |||||| |
Db   39399 ATGCCAACTAATTTTTTTT------TGTTTTTTTGAGATGGAGTTTGCTCTTGTTG 39348

Qy     643 CCCAGGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCA 702
           |||||||||||| ||||||||||||||| ||||| ||||| ||||||||||||||||| ||
Db   39347 CCAGGCTGGGGTGCCAAATGGTGCAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCA 39288

Qy     703 AGTTATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCCCTAG 762
           |||||||||||||||||||| ||| ||||||||||||||||||||| |||||||||| ||
Db   39287 AGCAATTCTCCTGCCTCAACCCCCCAAGTAGCTGGGATTACAAGCGCATGCCATCACACC 39228

Qy     763 CTAATTTTTTGTATTTTAGTAGAGACGGGG--TTCACAGTGTTAGCCAGGATGGTCTTG 821
           |||||||||| ||| | |||||||||||||  ||||| ||||||||||||||||||||||
Db   39227 TGGCTAATTTGTATTGTTAGTAGAGATGGGGTTTCACTATATGTCAGGATGGTCTTGATCTTG 39168

Qy     822 ATCTCTGACCT--TGTGATCTCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 879
           |||| ||||||  ||| ||||||||| ||||||||||||||||||||||||||||||||||
Db   39167 AACTCCTGACCTCAGGAGATCCTGCCTCGGCCTCCCAGAGGGGCTGGGATTACAGGCGTG 39108

Qy     880 TGAGCCACCACGCCCGGCCTT--ATTTTTTAATTTTGTTTTGTTGTTTGAAATGGAATCTCACTC 937
           | ||| ||||| ||||||||| ...
```

```
Db   39107 TGAGCCACTGTGCCTCGGCCTGCCTGCCCAATTTTTT...                39048
Qy     938 TGTTACCCAGGCTGGAGTGCAGTGGCAATGGCCAAATCTCGGCTCACTGCA... 997
Db   39047 TCTTTCCCAGGCTGGAGTGGAGGTGGAATGGCACCAC...                38988
Qy    1058 ACACCCGCTAATTTTTTGTATTTTTTTTAGTAGAGACGG...              1117
Db   38987 GTTCAAGCGATTCTCCTGCCTCAGTCTCCCAAGTAGCTGGGACTAC...      38928
Qy     998 GCTCAAGCGGATTCTCCTGTCTCAGCCTCCCAAGCACACTG...           1057
Db   38927 ACGCCCGGCCTAATTTTTGTATTTTTTAGTAGAGACGG...              38868
Qy    1118 GTCTCAAAACTCCTGACCTCAGGTGATCCACCCGCCTCAGC...           1177
Db   38867 GTCTTGAACTCCTTAACTTCAGGTGATCACCCGGCCTCGGCCTCCC...      38808
Qy    1178 CAGGCGTGAGCCACCTCACCCAGCC---GGCTAATTTAGATAAAAAAAAAAT... 1233
Db   38807 CAGGCATGAGCCACCACCACCTGGCCTCGCCTCTGCCCAACTAATT...      38748
Qy    1234 GGGGGGGTCTTGCTANGTGTTGCCCAGGCTGGTCTCAAACTCCTGG...      1293
Db   38747 ACAGGAGTCTCACTATGTTGTTCCAGACTGGTCTTGAACTCCTGG...       38688
Qy    1294 AAAATGAGCCAC 1304
           ||| |||||||
Db   38687 GTCTCAGCCTG 38677
```

RESULT 7

ABL67668/c

```
ID   ABL67668 standard; DNA; 65608 BP.
XX
AC   ABL67668;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Oesophagus cancer related gene sequence SEQ ID NO:6005.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PD   13-DEC-2001.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   25-SEP-2000; 2000US-235280P.
PR   25-SEP-2000; 2000US-235637P.
PR   26-SEP-2000; 2000US-235673P.
PR   26-SEP-2000; 2000US-235768P.
PR   27-SEP-2000; 2000US-235711P.
```

KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.

```
PR    27-SEP-2000;    2000US-235720P.
PR    27-SEP-2000;    2000US-235840P.
PR    27-SEP-2000;    2000US-235863P.
PR    28-SEP-2000;    2000US-236028P.
PR    28-SEP-2000;    2000US-236032P.
PR    28-SEP-2000;    2000US-236033P.
PR    28-SEP-2000;    2000US-236034P.
PR    28-SEP-2000;    2000US-236109P.
PR    28-SEP-2000;    2000US-236111P.
PR    29-SEP-2000;    2000US-236842P.
PR    29-SEP-2000;    2000US-236891P.
PR    02-OCT-2000;    2000US-237172P.
PR    02-OCT-2000;    2000US-237173P.
PR    02-OCT-2000;    2000US-237278P.
PR    02-OCT-2000;    2000US-237294P.
PR    02-OCT-2000;    2000US-237295P.
PR    02-OCT-2000;    2000US-237316P.
PR    02-OCT-2000;    2000US-237425P.
PR    03-OCT-2000;    2000US-237598P.
PR    03-OCT-2000;    2000US-237604P.
PR    03-OCT-2000;    2000US-237606P.
PR    03-OCT-2000;    2000US-237608P.
PR    01-NOV-2000;    2000US-244867P.
PR    01-NOV-2000;    2000US-245084P.
XX
PA    (AVAL-)  AVALON PHARM.
XX
PI    Young PE,  Augustus M,  Carter KC,  Ebner R,  Endress G,  Horrigan S;
PI    Soppet DR,  Weaver Z;
XX
DR    WPI;  2002-188264/24.
XX
PT    Claim 1;  SEQ ID 6005;  44pp;  English.
XX
PS    Claim 1;  SEQ ID 6005;  44pp;  English.
XX

CC    The present invention describes a method (M1) for screening for an
CC    anti-neoplastic agent. The method involves exposing cells to a chemical
CC    agent to be tested for anti-neoplastic activity, determining a change in
CC    expression of at least one gene (I) of a signature gene set, where (I)
CC    comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC    to ABL70110), or is at least 95% identical to (S), where a change in
CC    expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC    activity and can be used in gene therapy. M1 can be used for screening
CC    an anti-neoplastic agent, and can be used for producing a product which
CC    is the data collected with respect to the anti-neoplastic agent as a
CC    result of M1, and the data is sufficient to convey the chemical
CC    structure and/or properties of the agent. M1 can be used in the
CC    treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC    oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC    adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC    infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC    carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ    Sequence 65608 BP;  15987 A;  16979 C;  16761 G;  15878 T;  3 other;
```

```
Query Match     40.8%;  Score 587.8;  DB 24;  Length 65608;
Best Local Similarity   70.68;  Pred. No. 5.4e-146;
Matches  940;  Conservative    0;  Mismatches  347;  Indels  44;  Gaps  10;
```

```
Qy      4  TTTTTTTTTGAGATGGAGGAGTTTTTCGGCTCTGTTCCCAGGCTGGAGTGCAATCT                 63
           ||||  ||| ||||||||||||||||||||||||||||||||||||||||||| ||||
Db   39993  TTTTGTTTTGAGATGGAGAGTTTT--GCTCTGTTGTTGCCCAGGCTGAAGTGCACGGAATCT       39935

Qy     64  CAGCTCACCGCACCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCG              123
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   39934  CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAGCAAATTCTCCTACCTCAGCCTCCCGAG       39875

Qy    124  TAGCTGGGATTACAGGCATGTGCCACCACGCTCGGCTAATTTTGTATTTTTTAGTAG              183
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   39874  TAGCTGGGATTACAGGCATGTGCCACCACGCTAATTTTTGTATTTTTTTTAGTAG--AGT       39819
```

```
Qy    184  AGATGGAGTTTCTCACTGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATCCGCTT                    243
           |||||| |||||||| |||||||||||||| ||||||||||||| ||||||||||||| |
Db   39818  AGACGGGGTTTCTCCATTGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCGCC       39759

Qy    244  CCGTCTCCGCCTTCCCAAAGTGCTAGATACAGGACTGGTCTTGGTCTTTGACCTTGGTGATT                     303
           || |||||||||||||||| || |||||| ||||||| |||| |||||||||||||||| |
Db   39758  CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGCCACCGCGCCCGGCCTGATTA       39700

Qy    304  CTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGGCCAAGGCTGGTCTCA                          363
           || || || ||||||||| || |||||||| |||| |||||| |||||||||
Db   39699  AGTATGAAAGCAGTACCTCCGGAATATGCACGGCCCTCCAAAACCAGCCTCTGAAGAGACTC       39640

Qy    364  AGCAGTCCACCTGCCTCAGCCTCCCAGGTGCTGGGATTA                         403
           |||||| |||| ||||||||||||| ||||||||||||
Db   39639  TTCACTTGTTCCTCGGCTGGGGCCACCACGTTCCACCTAAAAACTGGGGCTGGTGTTCTCTT       39580

Qy    404  CAGGCGTGCAGCCGTGGGCCCTTGGCCTTTTATTTTTTATTTTTTTAAGACACAGGGTTTCA                     463
           ||||||||||||||||| |||||||||||||||||||||||||||
Db   39579  CCCCGCTTCCCATCCCTGGCGTCTTTTTATTTTTT-------------TTTTTGAGATGGAGT       39520

Qy    464  TCTTACCCAGGATGAAGTGCAGTGGTGTGAATCACAGCTCACTGCAGCCTTCAACTCCTGA                      523
           ||| ||||||||||||||| ||||| || |||| |||||||||||||||
Db   39519  TGTTGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCCTCCCGG       39460

Qy    524  GATC-AAGCAATCCTCCTGCCTCAGCCTCCCAGCCTTGAGTAGCTGGGACTATAGGTGCCCAC                     582
           |||  || ||||| ||||||||||||||||||| |||||||||||||||||||| ||||
Db   39459  GCTCAAAGGAATCCTCCACCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCACCACACT       39400

Qy    583  ACACCTGGCCTAATTTTTTTTTTTTTTTAATTTTTTTTTGAGGAGTCTCACTCTGTCA                       642
           ||||||||||||| ||||||||||||||||||||||||| |||||||||| |||
Db   39399  ATGCCCAACTAATTTTTTTTTTTT---------TTTTTTTGAGATGGAGTTCGCTCTTGTTG       39348

Qy    643  CCCAGGCTGGAGTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCA                      702
           ||||||||||||||||||||||| |||||| |||||||||||||||||||||||||
Db   39347  CCCAGGCTGGGGTGCAATGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCA       39288

Qy    703  AGTTATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCCCACCACGCCCTAG                      762
           || ||||||||| |||||||| |||||||||||||| |||| ||||||||||||||| |
Db   39287  AGCAATTCTCCTGCCTCAACCCCCAAGTAGCTGGGATTACAGGCATGTGCCACCATGCCCAC       39228

Qy    763  CTAATTTTTTGTATTTTTAGTAGAGACAGGGG-TTCACCATGTTGGCCAGGTTGATCTTG                       821
           ||||||||| |||||| |||||||||||||   |||||||||||| ||||||||||||
Db   39227  TGGCTAATTTGTTTGTAGAGATGAGGGTTTCACTATGTGGTCAGGCTGGTCTCGAACTC       39168

Qy    822  ATCTCTGGACCT--TGTGATCTGCCGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCG                        879
           ||||||| |||  |||| ||| |||||||||||| |||||||||||||||||||
Db   39167  AACTCCTGACCTCAGGAGATCCCTCGCCTCGGCCTCCCAGAGGCTGGGATTACAGGTG       39108

Qy    880  TGAGCCACCACGCCCGGCCTT--ATTTTTTAATTTTGTTTTGTTTGAAATGGAATCTCACTC                     937
           ||||||||||||||||||||  |||||||| ||||||| ||||||||||||||||||
Db   39107  TGAGCCACTGTGCCTGGCCTCCCAATTTTTTTTTTTTTTTGGAGACAAAAGTCTGC       39048

Qy    938  TGTTACCCAGGCTGGAGTGCAAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGG                      997
           |||||||||||||||||||||| || |||||||||||||||||||||||||||||||||
Db   39047  TCTTTCCCAGGCTGGAGTGCAGTGGAATGCACCACCGCACCTCGGCTTCCCTTCGG       38988

Qy    998  GCTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGCTAGGGATTACAGGCGCACCCACCACC                      1057
           |||||| |||||||||| || |||||||||||| |||||||||||||||| |||||
Db   38987  GTTCAAGCGATTCTCCTGCCTCAGCCTAGGAGACTACAGGCGCATGTGCCACC       38928

Qy   1058  ACACCCCGGCTAATTTTTTGTATTTTTTCATTTAGAGACGGGGTTTCACCATTATTTTGTCAGGCTG             1117
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db   38927  ACGCCCGGCTAATTTTTTGTATTTTTTAGTAGAGAGCAGTTTCACCATGTTGGCCAGACTG       38868

Qy   1118  GTCTCAAACTCCTGACCTCAGGTGACCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTA             1177
           ||||| |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   38867  GTCTTGAACTCCTAACTTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTA       38808

Qy   1178  CAGGCGTGAGCCACCCACCCCAGCC-----GGCTAATTTAGATAAAAAAATATGTAGCAAT             1233
           |||||| ||||||||||||||||||     |||||| |||||||||||| |||||||
Db   38807  CAGGCATGAGCCACCACCACCTGGCCTCTGCCTCCCAAACTAATTATAAAAAAAATTTTTTGAGAG       38748
```

```
Qy    1234  GGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAAACTTCTGGCCTTCATGCAATCCTTCC  1293
                  ||| ||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38747  ACAGGATCTCACTATGTTGTCTCCAGACTGGTCTCTGAACTCCTGGGCTTAAGGGATCGTCCT  38688
```

```
Qy    1294  AAATGAGCCAC  1304
                  |  ||||||||
Db   38687  GTCTCAGCCTC  38677
```

```
Db   38687  GTCTCAGCCTC  38677
```

RESULT 8

ID   ABL66947   standard; DNA; 62944 BP.
XX
AC   ABL66947;
XX
DT   15-MAY-2002 (first entry)
XX
DE   Lung cancer related gene sequence SEQ ID NO:5284.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO20019462 9-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   22-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   26-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   27-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   28-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236034P.
PR   28-SEP-2000; 2000US-236109P.
PR   29-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-236891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   02-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PS   Claim 1; SEQ ID 5284; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical agent's
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, Papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 62944 BP; 15145 A; 17344 C; 15900 G; 14555 T; 0 other;
```

Query Match            36.4%;  Score 524.4;  DB 24;  Length 62944;
Best Local Similarity  71.9%;  Pred. No. 4.3e-129;
Matches   963;  Conservative   0;  Mismatches  321;  Indels  56;  Gaps  19;

```
Qy      1  TTTTTTTTTTTGAGAGTGGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA   60
                  |||||||| |||| ||| || ||||||||| | ||||||||| |||||||||||||||||
Db  59357  TATTTATTTTTGTGAGATGGAGTTTT-GCTCTTCTCGCCCAGGCTGGAGTGCAGTGGCGCAA  59415
```

```
Qy     61  TCTCAGCTCACCGCAACCTCCGCCTCCGGGTTCAAGCGGATTCTCCTGCCTCAGCCTCCC   120
                  |||| | |||| |||||||| ||||||| ||||| || ||||||||||| |||||||||
Db  59416  TCTCGGCTCACTGCAACCTCCGCCTCCGAGTTCAATAGATTCTCCTGCCTCAGCCT---  59472
```

```
Qy    121  CAGTAGCTGGGATTACAGGCATGCACCCACCGCTCGGCTAATTTTGTATTTTTTTAG   180
                  | ||||||| |||||||||  || || ||||||||| ||||||| |  |||||||||
Db  59473  TAGTAGGTGTGATTACAGGCATCTGCACCACCACGCCCAGCTAATTTT---TGTATTTTTAC  59529
```

```
Qy    181  TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC   240
                  ||||||||||||  |||||||| ||||||||||||||| |||||| |||| |||| |||
Db  59530  TAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGATCTTAGGTTATC  59589
```

```
Qy    241  CCTCCCGTCTCGGCCTCCCAAAGTGCTAG----ATACAGGACTGGCCACCATGCCCGGCCTC   296
                  |||||||||||||||||||||||||||||    || |||| ||||||||| ||| ||||||
Db  59590  TGCCTCACCTCGGCCTCCCAAAGTGCTGGGATTACACACGTGCCACCGCGAGCCCAGCCTC  59649
```

```
Qy    297  TGCCTGGCTAAT-TTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCC   355
                  |||||||||  | |||| || |||| || ||||||||||| || |||| |||||||||||
Db  59650  TTTTTTCTTTTTCTTTTTCTTTTTTCTGAGGCAGGTCTCGCCACCCAGGCTGGAGTGC  59709
```

```
Qy    356  TG----AGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTG   411
                  ||    | ||||||  |||| || ||||||||||||||||||||| |||||||| |||
Db  59710  AGCAGCAGGATCATPAGCTCACTGAGCTTCGAATCTCCGGTTCACAGTGATCCTCCCAGCT  59769
```

```
Qy    412  CAGCCGTGCCTGGCCTTTTATTTTTATTTTTTTTTTTAAGACACAGGTGTCCCACTTTACC   471
                  ||||| ||| | |  || | |||||||||||||||||||||||||||| | || ||||||
Db  59770  AATTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTGAGATGGA-GTCTCGGCTGTGTCGCC  59828
```

```
Qy    472  AGGATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTCAACTCCTGAGATCAAGC   531
                  || | |||||||| ||||||||||||||||||||||||||| | | | || || |||||
Db  59828
```

```
RESULT 9
ABL68262
ID   ABL68262 standard; DNA; 62944 BP.
XX
AC   ABL68262;
XX
DT   15-MAY-2002  (first entry)
XX
DE   Kidney cancer related gene sequence SEQ ID NO:6599.
XX
KW   Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW   stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW   cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW   gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200194629-A2.
XX
PD   13-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US10838.
XX
PR   05-JUN-2000; 2000US-209473P.
PR   05-JUN-2000; 2000US-209531P.
PR   18-SEP-2000; 2000US-233133P.
PR   18-SEP-2000; 2000US-233617P.
PR   20-SEP-2000; 2000US-234009P.
PR   20-SEP-2000; 2000US-234034P.
PR   20-SEP-2000; 2000US-234052P.
PR   20-SEP-2000; 2000US-234509P.
PR   22-SEP-2000; 2000US-234567P.
PR   25-SEP-2000; 2000US-234923P.
PR   25-SEP-2000; 2000US-234924P.
PR   25-SEP-2000; 2000US-235077P.
PR   25-SEP-2000; 2000US-235082P.
PR   25-SEP-2000; 2000US-235134P.
PR   25-SEP-2000; 2000US-235280P.
PR   26-SEP-2000; 2000US-235637P.
PR   26-SEP-2000; 2000US-235638P.
PR   27-SEP-2000; 2000US-235711P.
PR   27-SEP-2000; 2000US-235720P.
PR   27-SEP-2000; 2000US-235840P.
PR   28-SEP-2000; 2000US-235863P.
PR   28-SEP-2000; 2000US-236028P.
PR   28-SEP-2000; 2000US-236032P.
PR   28-SEP-2000; 2000US-236033P.
PR   28-SEP-2000; 2000US-236109P.
PR   28-SEP-2000; 2000US-236111P.
PR   29-SEP-2000; 2000US-236842P.
PR   29-SEP-2000; 2000US-236891P.
PR   02-OCT-2000; 2000US-237172P.
PR   02-OCT-2000; 2000US-237173P.
PR   02-OCT-2000; 2000US-237278P.
PR   02-OCT-2000; 2000US-237294P.
PR   02-OCT-2000; 2000US-237295P.
PR   02-OCT-2000; 2000US-237316P.
PR   03-OCT-2000; 2000US-237425P.
PR   03-OCT-2000; 2000US-237598P.
PR   03-OCT-2000; 2000US-237604P.
PR   03-OCT-2000; 2000US-237606P.
PR   03-OCT-2000; 2000US-237608P.
PR   01-NOV-2000; 2000US-244867P.
PR   01-NOV-2000; 2000US-245084P.
XX
PA   (AVAL-) AVALON PHARM.
XX
PI   Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI   Soppet DR, Weaver Z;
XX
DR   WPI; 2002-188264/24.
XX
PS   Claim 1; SEQ ID 6599; 44pp; English.
XX
CC   The present invention describes a method (M1) for screening for an
CC   anti-neoplastic agent. The method involves exposing cells to a chemical
CC   agent to be tested for anti-neoplastic activity, determining a change in
CC   expression of at least one gene (I) of a signature gene set, where (I)
CC   comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC   to ABL70110), or is at least 95% identical to (S), where a change in
CC   expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC   activity and can be used in gene therapy. M1 can be used for screening
```

```
Db  59829  AGACTGGAGTGCAGTGGCACAAATCTCAGCTCACTGCAACCTT-GCCTCCTGGATTCAAGC  59887
Qy    532  ATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGC  591

Db  59888  AATTCTCTGCCTCAGCCTCTTGA-GTAGCTAGGATTACAGGCGCTGCCACCACGCCCGGA  59946
Qy    592  TAATTTTTTA--TTTTTATTTATTTTTAATTTTTTTGAGACGGAGTCTCACTCTGTCACCAGGC  649

Db  59947  TAAATTTTTGGGTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTC-ACCCTGTCGCCAGGC  60005
Qy    650  TGGAGTGCAGTGGCGCAAATCTTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATT  709

Db  60006  TGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAAGCTCCGGGTTCATGCCATT  60065
Qy    710  CTCCTGCCC------CAGCTCCTGAGTAGCTGGGACTACAGGCGC---CCACCACGCC  759

Db  60066  CTCCCGCCTCCCAAGTAGCTGGGACTACAGGCACCGCGCCTCCGGGTTCATGCCATT  60125
Qy    760  TAGCTAATT---TTTTTGTAATTTTTTAGTAGAGATGGGG-TTCACCATGTTGCCCAGGCTG  815

Db  60126  CAGTTAATTTTTTGTTTTTTTGTAATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG  60185
Qy    816  ATCTTGATCTCTGGACCTTGTGATCTGCCCGCCTCCCAAAGTGCTGGGATTACA  875

Db  60186  GTCTCAAATCTCCTGACCTGGTGATCCGCCCGCCTCACCCTCCCAAAGTGCTAGGATTACA  60245
Qy    876  GGCGTGAGCCACCACGCCCGGCCTTATT----------TTTTAATTTTTGTTTGTTTGAAAT  925

Db  60246  GGCTGAGGGACCGCCCGCGCCTGGCCGTGTTTTTTTGTTTTGTTTTGTTTTTTTTTGGACG  60305
Qy    926  GGAAATCTCACTCTGTTACCCAGGCTGGAGTGCAGTGGCGCAAATCTCGGCTCACTGCAACC  985

Db  60306  GAAATTTTGCTCTTGTTGTCCCAAGCTGGAGTGGCAATCTCAGCTCACTGCAACC  60365
Qy    986  TCTGCCTCCCGGGGCTCAAGCGGATTCTCCTGTCTCTCAGCCTCCCAAGCAGCTGGGATTACGG  1045

Db  60366  TCTGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACCG  60425
Qy   1046  GCACCTGCCACCACACCCGGCTAA-TTTTTGTTATTTTCATTAGAGGCGGGGTTTCACCAT  1104

Db  60426  GCGTGTGAGGCGACCGCCCGTGTTTTAGTTATTTTTAGTAGAAATGGGGTTTCACCAT  60485
Qy   1105  ATTTTGTCAGGCTGGTCTCGAAACTCCTCGACCTCAGGTGATCACCCGCCTCGGCCTCCCAA  1164

Db  60486  ATTGGTCAGGCTGGTCTCAAACTCTTGACCTC--GTGATCCACCCGCCTCGGCCTCCCAA  60543
Qy   1165  AGTGCTGGGATTACAGGCGTGAGCCACCACCCAGCCGGCTAATTTTAGATAAAAAAATA  1224

Db  60544  AGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCC-------------------T  60596
Qy   1225  TGTAGCAAATGGGGGGGGTCTTGCTATGTTGCCCAGGCTGTCTCAAAACTTCTGGCTTCATGC  1284

Db  60597  TTAATAGAGATGAGGTCACACTGTGTTGCCCAGGACAGTCTCGAACTCCTGAACTTCACAC  60656
Qy   1285  AATCCTTCCAAATGAGCCAC  1304

Db  60657  AATCTGCCTCGCCTCAGCCTC  60676
```

```
CC   an anti-neoplastic agent, and can be used for producing a product which
CC   is the data collected with respect to the anti-neoplastic agent as a
CC   result of M1, and the data is sufficient to convey the chemical
CC   structure and/or properties of the agent. M1 can be used in the
CC   treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC   oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC   adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC   infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC   carcinoma, papillary carcinoma and Wilm's tumour.
XX
SQ   Sequence 62944 BP; 15145 A; 17344 C; 15900 G; 14555 T; 0 other;

Query Match             36.4%;  Score 524.4;  DB 24;  Length 62944;
Best Local Similarity   71.9%;  Pred. No. 4.3e-129;
Matches   963;  Conservative    0;  Mismatches  321;  Indels   56;  Gaps   19;

Qy      1 TTTTTTTTTTTTTTGAGAGGAGTGTTTTTTGGCTCTTGTTTGCCAGGTGGACTGCAAATGGGCGAA 60
          |||||   ||||||||||||||   |||||||||| ||||||||||||||| ||| |||||||| 
Db  59357 TATTTATTTTTTGAGAGGAGTTTT-CCTCTTCTCGCCCAGCCTGGGAGTGCAGTGGGCGCAA 59415

Qy     61 TCTCAGCTCACCGCAACCTCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          ||||  ||||| |||||||||| |||    |||||||||||||||||||||| 
Db  59416 TCTCGTCTCACTGCAACCTCCGCCTCCTGAGTTCAAATAGATTCTCCTGCCTCAGCCTC--- 59472

Qy    121 CAGTAGCTGGGATTACAGGCGATGTGCACCACGCTCGGCTAATTTTTGTATTTTTTTTTAG 180
          |||||| || ||||||| |||| | | |||||| || ||||||||   |||||||||||  
Db  59473 TAGTAGGTGTGAATTACAGCGATCTCGCCACGCTCAAATTTT---TGTTATTTTTTAC 59529

Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCACGACCTCAGATGATC 240
          ||||||  ||||||| ||||||||| |||||||||||||||||| |||  ||||| ||| ||
Db  59530 TAGAGACAGGGTTTCACCAATGTTGGCCAGGCTGGTCTTGAACTCCTGATCTTAGGTTATC 59589

Qy    241 CCTCCGTGCTCGGGCCTCCCAAAGTGCTAG---ATACAGAGACTGGCCACCACTACACCTGG 296
          | || |||| ||||||||||||||||||   | || ||| |||||| |||||| || |||| 
Db  59590 TGCCTACCCTGCGCCCTCCCAAAGTGCTGGGATTACACACGTGAGCGCCACCGCGCCCAGCCTC 59649

Qy    297 TGCCTGGCTAAT-TTTTGTGGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCC 355
          |    | ||||| |||| |   |     | ||||  ||| | ||| ||| ||||||||||| 
Db  59650 TTTTTTCTTTTTTTTTTCTTTTTTTTCGAGGCAGGGTCTCGCTCTGTCACCCAGGCTGGAGTGC 59709

Qy    356 TG---AGCTCAAGCAGTCCACCTCACCTCCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTG 411
          |    ||| |   ||| | || | |  || | |||||||||||||| |||||||||| || || 
Db  59710 AGCAGCAGGATCATAGCTCACTGCAGCTTCGAACTCCCGGTTCAAGTGATTCTCCCAGCT 59769

Qy    412 CAGCCGTGCCTGGCCTTTTTTATTTTTATTTTTTTAAGACACAGGTGTCCCACTCTTACCC 471
          |   | |   ||| ||||||| |||||||||||||| | |||| || ||| ||| ||||| 
Db  59770 AATTTTTTTTTATTTTTATTTTTTTTTTTGAGATGGA-GTCTGGCTCTGTGTCGCCC 59828

Qy    472 AGGATGAAGTGCAAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGC 531
          |||| ||| ||| || || | || | ||||||||||| ||||||||||||| || ||||| 
Db  59829 AGACTGGAGTGCAGTGGCGACAAATCTCAGCTCACTGCAACCTT-GCCTCCTGGATTCAAGC 59887

Qy    532 ATCCTCCTGCCTCAGCCTCCCGAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGC 591
          | | ||| |||| ||||||| || |||||| || | || |||| |||||| || ||||   
Db  59888 AATTCTCTGCCCTCAGCCTCTGA-GTAGCTAGGATTACAGGCGTGCCACCACGCCCGGA 59946

Qy    592 TAATTTTTTA--TTTTTATTTTTAATTTTTTGAGACAGAGTCTCAACTCGTCACCCAGGC 649
          ||| |||| |  ||||||||||| ||||||||| ||||||||| | || || |||| || 
Db  59947 TAAATTTTGGGTTTTTTTTTTTTTTTTTGACACAGAGTCTC-ACCCTGTCGCCCTAGCC 60005

Qy    650 TGGAGTGCAGTGGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATT 709
          |||||||||||||  | ||||| ||||||||||||| || |||||||||||||||  ||| 
Db  60006 TGGAGTGCAGTGGTGCGGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCAAGCCATT 60065

Qy    710 CTCCTGCCC-----CAGCCTCCTGAGTAGCTGGGACTACAGGCGC--CCACCACGCC 759
          |||| | ||     | ||||||| ||||||||||||||||||| |   |||||| || 
Db  60066 CTCCCGGCCTCCCAAGTAGCTGGGACTACAGGCACCACCATGCC 60125

Qy    760 TAGCTAATT---TTTTGTTATTTTTAGTAGAGATAGGGG-TTCACCATGTTCGCCAGGTG 815
          ||| ||||    |||| | |||||||||||||| ||||| ||||||| ||| |||||| | 
Db  60126 CAGTTAATTTTTGTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATG 60185

Qy    816 ATCTTGATCTCTGGACCTTGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACA 875
          | || |||||||||| || |||||||||| |||||||||||||| | |||||||||||| 
Db  60186 GTCTCAAATCTCCTCGACCTGGTGATCCGCCCGCCTCGGCCTCCCAACCTCCCAAAGTGCTAGGATTACA 60245

Qy    876 GGGCGTGAGCCACCACGCCCGGCCTTATT------- 925
          || | |||| |||||||||| ||||  |        
Db  60246 GGCTTGAGGCACCACGCCCTGGCCTGGTTTTTTTTGTTCGTTTTTTTTTTGGACG 60305

Qy    926 GGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAATGGGCCAAAAATCTCGGCTCACTGCAACC 985
          | ||| | ||||||| || ||| ||||||||| || || |||   |||| ||||||||||||| 
Db  60306 GAATTTTGCTCTGTTGTTGCCCAAGCTGGAGTGCGAATAGGTGGCCATCTCAGCTCACTGCAACC 60365

Qy    986 TCTGCCTCCCGGGCTCAAGCGAGATTCTCCTGCCTCAGCCTCCCCAAGCAGCGTGGGATTACGG 1045
          ||||||||| || ||||||||  ||||||| ||||||||||||  |||||||||||||||| | 
Db  60366 TCTGCCTCCCAGTTCAAGCGATTCTCCCGTCCTCAGCCTCCTGAGTAGCTGCGATTACCG 60425

Qy   1046 GCACCTGCCACCACACCCGGCTAA-TTTTGTATTTTCATTAGAGGCGGGGTTTCACCAT 1104
          || |  |||||||| ||||||||| ||||||||||  | ||||   |||||||||||||
Db  60426 GCGTGTGCCACCATGCCCGGCTAATTTTTGTATTTTTAGTAGAAATGGGGTTTCACCAT 60485

Qy   1105 ATTTGTCAGGCGTGGTCTCAAACTCCTGACCTCAGGTGACCCACCTGCCTCAGCCTTCCAA 1164
          |||  |||||| ||||||||||||| |||||||  ||||||||| |||||| |||||||||
Db  60486 ATTGGTCAGGCTGGTCTCAAACTCTTGACCTC--GTGATCCACCCGCCTCGGCCTTCCAA 60543

Qy   1165 AGTGCTGGGATTACAGGCGTGAGCCACCACGCCCAGCCTGAATTTAGATAAAAAAATA 1224
          |||||||||||||||||| |  | |||| || |||||| || |||||||      | 
Db  60544 AGTGCTGGGATTACAGGCATAAAACCACTGCGCCTGGCCTAATTTAGATA------T 60596

Qy   1225 TGTAGCAAATGGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAAACTTCTGGCTTCATGC 1284
          | | | || | | |||||  ||| |||||||||| | ||||||  ||||||||| | |   |
Db  60597 TTAATAGAGATGAGGTCACACTGTGTTGCCCAGGACAGTCTCGGAACTCCTGAACTCCTGAATTCAAC 60656

Qy   1285 AATCCTTCCAAATGAGCCAC 1304
          ||||  | |  |||||||| 
Db  60657 AATCTGCCTCGCCTCAGCCTC 60676
```

RESULT 10
AAK69279/c
```
ID   AAK69279 standard; DNA; 33747 BP.
XX
AC   AAK69279;
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24091.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
```

PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235634.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 24091; 307lpp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially

```
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 33747 BP; 9625 A; 7893 C; 7691 G; 8538 T; 0 other;

Query Match             35.9%;  Score 517.2;  DB 22;  Length 33747;
Best Local Similarity   68.8%;  Pred. No. 2.8e-127;
Matches  844;  Conservative   0;  Mismatches  363;  Indels  19;  Gaps  9;

Qy      1 TTTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
          ||||||||||||| ||||| ||||||||||     ||||||||||| |||||||||||||||
Db  33207 TTTTTTTTTTTTTGAGATGGGAGAGTCT--CGCTCGTCGCCCACCAGGTTGGAGTGCAGTGATGAGTGA 33150

Qy     61 TCTCAGCTCACCCGCAACCTCCGCCTCCCGGGTTCAAGCGAATTCTCCTGCCTCAGCCTCCC 120
          |||||||||||||||||||||||||||| |||||||||| ||||||||||||||||||||||
Db  33149 TCTCAGCTCACTGCAACCTCTGCCTCCCAGGTTTAAGCGAATTCTCCTGACTTAGCCTCCC 33090

Qy    121 CAGTAGCTGGGATTACAGACGCATGTGCCACCCACGCTCGGCTAATTTTGTATTTTTTTTAG 180
          ||||||||||||||||||||||  |||||||||||||||||||||||||| ||||||||||
Db  33089 TAGAAGCTGGGATTATAGGCGCTGCCACCACATGCCCTGGCTAATTTTT---TGTATTTTTAG 33033

Qy    181 TAGAGAATGGAGTTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
          |||||||||| |||||||||||| | ||||||||| |||||||||||||||| ||||||||
Db  33032 CAGAGACGGGGTTTTGCCATGTTGGCCAGGCTGTCTCAAACTCCTGAACTCAGGTGATC 32973

Qy    241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCGCGCCCGGCCTGCC 300
          ||||| ||||||||||||||||||||| |||||||||||| ||||||||| ||||||||||
Db  32972 TGCCTGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCGCCCGGCCTGCC 32913

Qy    301 TGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCC-- 355
          ||||||||||||||||| |||||||||||||||||| ||||||||||||| |||||
Db  32912 TTTTTTTTATTTATTTTGTTGAGACAGGAATCTTGCTCTTTTGTCCAGGCTGGAGTGCAGTGG 32853

Qy    356 TGAGCTCAAGCAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGC 415
          | ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  32852 TATGATCACAGCTCATTGCAGCCTCCAACTCCTGGACTCAGGCAATCCTCCCACCTTAGC 32793

Qy    416 CGTGCCTGGCCGCTTTTTATTTTTATTTTTAATTTTTTGAGACACACAGGTGTCCCACTCTT 475
          |||||||||||||||||||||| ||||||||||||| |||||||||||||||||||
Db  32792 TGCCTGAGTAGCTGGGACTACAAGCGTTGCGCCACCATGTCTGGCTAATTTTTGTATTTTTT 32733

Qy    476 TGAAGTGCAGTGGTGTGATCA--CAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGC-A 532
          ||| |||||||||||||||||  |||||||||||| |||||||||||| ||||||||
Db  32732 TGTAGAGATGGGGTTTTGCCATGTTGGTCAGGCTGGTCTCGAACTCCTGGGCTCAAGCAA 32673

Qy    533 TCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACCACACCTGGCT 592
          ||||||||||||||||||||||| ||||||||| ||||| |||||||||||||| |||
Db  32672 TCCTCCCGCCTTGGTTACCAAAGTGTGGGATTATAGGCGTGAGCCAPTGTGCCCAGACC 32613

Qy    593 AATTTTTATTTTTATTTTTTGAGACGAGTCTCAACTCTGTCACCCAGGCTGG 652
          ||||||||||||||| |||| |||||||||||||||||||||||||||||||
Db  32612 TATACAT-TGGCTTTTTTTTTTTAGATGGAGTCTC-ACTCTGTTGGCCAGGTCTG 32555

Qy    653 AGTCAGTGGCGCAAATCTTGGCTCACTGCAACCTCCGCCTCCCTGCCTCCGGGTTCAA 712
          ||||||||||| |||||||| |||||||||||||||||||||||||||| ||||||||
Db  32554 AGTGCAGTGGCGCAATCTTGGCTTACTGTGTAACCTCTCTCCCGGGTTCAAGCGATTCTC 32495

Qy    713 CTGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGC---CCACCACGCCTAGCTAATTT 769
          ||||||||||||||||||| |||||||||||||||||   |||||| |||||||||||||
Db  32494 CTGCCTCAGCCTCCTAAGTAGCTGGGACTAPAGGCGCATCGCCACCATGCCCGGCCAATTT 32435

Qy    770 TTTTGTATTTTAGTAGAGATGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCTTGG 829
          ||||||||||||||||| |||||| |||||||||||| ||||||||||||||||||||||
Db  32434 TTGTAATTTTTAGTAGAGAATGGGTTTTACCATGTTGGCCAGGCTGCTCTCGAACTTCTG 32375

Qy    830 ACCTTGTGATCTGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTATAGGCGTGAGCCACCA 889
          ||||||| |||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  32374 ACCTTGTGATCGTCTGCCTTGGCCTCCCAAAATGCTGGGATTACAGAGCACTG 32315

Qy    890 CGCCCGGCCTTATTTTTTAATTTTTGTTTGTTTGAAATGGGAATTCTCACTCTGTTACCCAGGC 949
          | |||||||| |||||| |||||| |||||||||  ||||||||| ||||||||||||||||||
```

```
Db  32314 CACCCGCCCGACTTTT-TTTTTCTTGGAGACGGAGTCTCACTCACTTTGTCGCCCAGGC 32256

Qy    950 TGGAGTGCAATGGCCAAAATCGGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATT 1009
          ||||||||||| |||||||||||||||||||||| ||||||||||| ||||||||||||
Db  32255 TGGAGTACAGTGGCGCATGAGTCTCAGCTCACTGCATCACCTCCTCCATCCTCCTGGGTTCATT 32196

Qy   1010 CTCCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGCGCACCACCACACCCCCGTTCAGC 1069
          ||||| ||||||||||||||||||| |||||| ||||||| |||| |||| |||||||
Db  32195 TTCCTGCCTCAGCCTCCCGAGTAGCTGCGATTACAGGCGTATGCCACCATACCCGTGGAGC 32136

Qy   1070 TTTTTTGTATTTCATTAGAGGGCGGGGTTTCACCATATTGTCAGGGTGGTCTCAAACTCC 1129
          |||||| |||| |||||| ||| |||||||||||||| ||||||||||||||| |||||
Db  32135 TTTTTGTTTTTAGTGAGATGGGGTTTTTACCATGTTGGTAAGCTGGTCTTCCAACTCC 32076

Qy   1130 TGACCTCAGGTGACCCACCTGCCTCAGCCTCCAAAGTGCTGGGATTACAGGCGTGAGCC 1189
          || |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  32075 TGACCTTAAGTGATCTGCCCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCC 32016

Qy   1190 ACCTCACCCAGCCGCGGCTAATTTAGAT 1215
          || ||||||||||||||||| || |||
Db  32015 ACTGTGCTCGGCCCCATTGACTTTTTAT 31990
```

PR   14-AUG-2000;   2000US-0225759.
PR   18-AUG-2000;   2000US-0226279.
PR   22-AUG-2000;   2000US-0226681.
PR   22-AUG-2000;   2000US-0226868.
PR   22-AUG-2000;   2000US-0227182.
PR   23-AUG-2000;   2000US-0227009.
PR   30-AUG-2000;   2000US-0228924.
PR   01-SEP-2000;   2000US-0229287.
PR   01-SEP-2000;   2000US-0229343.
PR   01-SEP-2000;   2000US-0229344.
PR   01-SEP-2000;   2000US-0229345.
PR   05-SEP-2000;   2000US-0229509.
PR   05-SEP-2000;   2000US-0229513.
PR   06-SEP-2000;   2000US-0304437.
PR   06-SEP-2000;   2000US-0230438.
PR   08-SEP-2000;   2000US-0231242.
PR   08-SEP-2000;   2000US-0231243.
PR   08-SEP-2000;   2000US-0231244.
PR   08-SEP-2000;   2000US-0231413.
PR   08-SEP-2000;   2000US-0231414.
PR   08-SEP-2000;   2000US-0232080.
PR   08-SEP-2000;   2000US-0232081.
PR   12-SEP-2000;   2000US-0231968.
PR   14-SEP-2000;   2000US-0232397.
PR   14-SEP-2000;   2000US-0232398.
PR   14-SEP-2000;   2000US-0232399.
PR   14-SEP-2000;   2000US-0233400.
PR   14-SEP-2000;   2000US-0232401.
PR   14-SEP-2000;   2000US-0233063.
PR   14-SEP-2000;   2000US-0233064.
PR   21-SEP-2000;   2000US-0233065.
PR   21-SEP-2000;   2000US-0234223.
PR   25-SEP-2000;   2000US-0234274.
PR   25-SEP-2000;   2000US-0234997.
PR   26-SEP-2000;   2000US-0234998.
PR   27-SEP-2000;   2000US-0235834.
PR   27-SEP-2000;   2000US-0235804.
PR   29-SEP-2000;   2000US-0236327.
PR   29-SEP-2000;   2000US-0236357.
PR   29-SEP-2000;   2000US-0236368.
PR   29-SEP-2000;   2000US-0236369.
PR   29-SEP-2000;   2000US-0236370.
PR   02-OCT-2000;   2000US-0236802.
PR   02-OCT-2000;   2000US-0237037.
PR   02-OCT-2000;   2000US-0237038.
PR   02-OCT-2000;   2000US-0237039.
PR   02-OCT-2000;   2000US-0237040.
PR   13-OCT-2000;   2000US-0239935.
PR   13-OCT-2000;   2000US-0239937.
PR   01-NOV-2000;   2000US-0240960.
PR   20-OCT-2000;   2000US-0241221.
PR   20-OCT-2000;   2000US-0241785.
PR   20-OCT-2000;   2000US-0241786.
PR   20-OCT-2000;   2000US-0241787.
PR   20-OCT-2000;   2000US-0241808.
PR   20-OCT-2000;   2000US-0241809.
PR   20-OCT-2000;   2000US-0241826.
PR   20-OCT-2000;   2000US-0244617.
PR   08-NOV-2000;   2000US-0246474.
PR   08-NOV-2000;   2000US-0246475.
PR   08-NOV-2000;   2000US-0246476.
PR   08-NOV-2000;   2000US-0246477.
PR   08-NOV-2000;   2000US-0246478.
PR   08-NOV-2000;   2000US-0246523.
PR   08-NOV-2000;   2000US-0246524.
PR   08-NOV-2000;   2000US-0246525.
PR   08-NOV-2000;   2000US-0246526.
PR   08-NOV-2000;   2000US-0246527.
PR   08-NOV-2000;   2000US-0246528.
PR   08-NOV-2000;   2000US-0246532.
PR   08-NOV-2000;   2000US-0246533.
PR   08-NOV-2000;   2000US-0246609.
PR   08-NOV-2000;   2000US-0246610.

PR   08-NOV-2000;   2000US-0246611.
PR   08-NOV-2000;   2000US-0246613.
PR   17-NOV-2000;   2000US-0249207.
PR   17-NOV-2000;   2000US-0249208.
PR   17-NOV-2000;   2000US-0249209.
PR   17-NOV-2000;   2000US-0249210.
PR   17-NOV-2000;   2000US-0249211.
PR   17-NOV-2000;   2000US-0249212.
PR   17-NOV-2000;   2000US-0249213.
PR   17-NOV-2000;   2000US-0249214.
PR   17-NOV-2000;   2000US-0249215.
PR   17-NOV-2000;   2000US-0249216.
PR   17-NOV-2000;   2000US-0249217.
PR   17-NOV-2000;   2000US-0249218.
PR   17-NOV-2000;   2000US-0249244.
PR   17-NOV-2000;   2000US-0249245.
PR   17-NOV-2000;   2000US-0249264.
PR   17-NOV-2000;   2000US-0249265.
PR   17-NOV-2000;   2000US-0249297.
PR   17-NOV-2000;   2000US-0249299.
PR   17-NOV-2000;   2000US-0249300.
PR   01-DEC-2000;   2000US-0250160.
PR   01-DEC-2000;   2000US-0250391.
PR   05-DEC-2000;   2000US-0251030.
PR   05-DEC-2000;   2000US-0251988.
PR   05-DEC-2000;   2000US-0256719.
PR   06-DEC-2000;   2000US-0251479.
PR   08-DEC-2000;   2000US-0251856.
PR   08-DEC-2000;   2000US-0251868.
PR   08-DEC-2000;   2000US-0251869.
PR   08-DEC-2000;   2000US-0251989.
PR   08-DEC-2000;   2000US-0251990.
PR   11-DEC-2000;   2000US-0254097.
PR   05-JAN-2001;   2001US-0259678.

XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC, Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
PS   Disclosure; SEQ ID NO 27905; 307pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 33747 BP; 8538 A; 7691 C; 7893 G; 9625 T; 0 other;

Qy   1 TTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGCGCAA   60
           |||||||||||||||||||||||||||||||| | ||||||| |||||| || | |
Query Match   35.9%; Score 517.2; DB 22; Length 33747;
Best Local Similarity   68.8%; Pred. No. 2.8e-127;
Matches   844; Conservative   0; Mismatches 363; Indels   19; Gaps   9;

```
Db    541 TTTTTTTTTTTTTTGAGATGGAGTGCT--CGCTCTGTCGCCCCAGGTTTGGAGTGCAGTGATGATGA  598
           |||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Qy     61 TCTCAGACTCACCGACAACCTCCGCGTCTCCGGTTCAAACGGATTTCTCTGCCTCGAGCCTCCC  120
Db    599 TCTCAACCTCACTGCAACCTCTGCCTCCCGGTTTAAGCAGGATCCTCGACTTAGCCTCC  658
Qy    121 CAGTAGCTACTTGGGATTACAGACGCATTGGACCACGCTCGGGCTAATTTGTTGTATTTTTTTTAG  180
           |||||||||||||||||||||||||||||||||||||||||||| |  ||||||||||||||
Db    659 TAGAAGCTGGGATTATAGGCGTGCACCACCATGCCTGGCTAATTT---TGTATTTTTAG  715
Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC  240
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    716 CAGAGACAGGGGTTTTTGGCCATGTTGGCCAGGCTGGTCTCAAACTCCTGAACTCAGGTGATC  775
Qy    241 CCTCCGTCTCCTGGGCCTCCCCAAAGTGTCTAGATACAGGTGTGAGCCACCATGCCCGGCCTA  300
           |||||||||||||||| |||||| ||||||||||||||||||||||||||||||||||
Db    776 TGCCTGCCTTGGGCCTCCCCATGTTGCCCAGGCTGGTCTCGAACTCCTGGCCTCAGGTGATC  835
Qy    301 TGGCCTAAATTTTTGTCGGTAGAAACAGGGTTTCACTGCGTTGTCCAGGCTGGAGTGCAGTGG  355
           |||||| |||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db    836 TTTTTTATTATTTTTTGTTGAGACAGGGTTTCTCTTTGTCCAGGCTGGAGTGCAGTGG  895
Qy    356 TGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGC  415
           |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db    896 TATGATCACAGCTCATTGCAGCCTCCAACTCCTGGACTCAGCGAATCCTCCCACCTTAGC  955
Qy    416 CGTGCCCTGGCCTTTTTATTTTTATTTTTTTTAAGACACAGGTGTGTCCACTCTTACCCAGGA  475
           || ||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db    956 TGCCTGAGTAGCTGGGGACTACAAGCTTGCGCCACCATGTCTGGCTAATTTTTGTATTTTT  1015
Qy    476 TGAAGTGCAGTGGTGTGATCA--CAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGC-A  532
           |||||| ||||||||||||||  |||||||||||||||||||||||||||||||||||| |
Db   1016 TGTAGAGATGGGGTTTTGCCATGTTGGTCTGTTGAACTCCTGGGCTCAAGCAA  1075
Qy    533 TCCTCCTCCCTCAGCCTCCCCAAGTAGCTGGGACCCAAAAGACATGCCACCACCTGGCT  592
           ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db   1076 TCCTCCCGCCTTGGTTTACCAAAGTGTTGGGATTATAGGCGTGAGCCATTGTTGCCACC  1135
Qy    593 AATTTTTATTTTTTATTTTTTAATTTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGG  652
           ||||||||| |||| |||||| ||||||||||||||||| |||||||||||||||||||||
Db   1136 TATACAT-TGGCTTTTTTTTTTTTTTTTTTTTAGATGGGAGTCTC-ACTCTGTTGGCCAGGGTGG  1193
Qy    653 AGTGCAGTGGCGTGGGCCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTC  712
           ||||||||||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||
Db   1194 AGTGCAGTAGCGTGGGCCAAATCTTGGCTTACTGTAACCTCTCCCGGGGATTCTC  1253
Qy    713 CTGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGC---CCACCACGCCCAGCTAATTTT  769
           ||||  ||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db   1254 CTGCCTCAGCCTCCTAGGTAGCTGGGACTATAGGCGCATGCCACCACACGGCCAATTT  1313
Qy    770 TTTTTTGTATTTTAGTAGAGATGGGGTTTCACCATGTTGATCTTGATCTCTGG  829
           |||| ||| ||||||||||||||||||||||||||| ||||||||||||||||
Db   1314 TTGTATTTTTTAGTAGAGATGCGTTTCACCATATTGGCCAGGCTGGTCTCGAACTTCTG  1373
Qy    830 ACCTTGTGATCTCGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCA  889
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1374 ACTTTGGATCTGTCTGCCTCCCTAGGTGCTGGGATTACAGGCGTGAGCCACCACTG  1433
Qy    890 CGCCCGGCCGCTTATTTTTTTTTTTTGTTTTTTTGTTTGAAATGGAAATCTCACTCTGTTACCCAGGC  949
           || |||||||||| ||||||||||||||||||||| ||||||||||||||||||||||||
Db   1434 CACCCGGCCGACTTTT-TTTTTTCTTCACTCACTTTGTGCCCAGGC  1492
Qy    950 TGGAGTGCAATGGCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCGAAGCGGATT  1009
           |||||||||||||||||||||||| ||||||||||||||||||||||||| |||||||||
Db   1493 TGGAGTACAGTGGCATGATCTCAGCTCACTATAACCTCCATCTCCTGGGTTCACGTCATT  1552
Qy   1010 CTCCTGTCTCAGCCTCCCAAGCAGCGTGGGATTACGGGCGCCACCACCACCCCGCTAA  1069
           ||||||||||||||||||| ||||||||||||| |||||| ||||||||||||||
Db   1553 TTCCTGCCTCAGCCTCCCATCCCCAGTAGCTGGGATTACAGGCGCCCACCACTACCCGGCTAA  1612
Qy   1070 TTTTTTGTATTTTTCATTAGAGGGCGGGGTTTTCACCATGTTGGTCTCAAACTCC  1129
           |||||| ||| ||||| ||||||||||| |||||||||||||||||||||||||||
Db   1613 TTTTTTGTGTTTTTAGTGAGATGGGGTTTTTACCATGTTGGGTAAGCTGGTCTCCAACTCC  1672
```

```
Qy   1130 TGACCTCAGGTGACCCACCTGTCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCC  1189
           |||||| ||| ||  ||||||||| | |||||||||||||||||||||||||||||||||
Db   1673 TGACCTTAAGTGATCTGCCCCGCCTCCAGCCTCCCAAAGTGCTGGGATTACAGGCCATGAGACGCC  1732
Qy   1190 ACCTCACCCCAGCGCGGCTAATTGTAGAT  1215
           ||  ||||| | |||  ||||||| ||||
Db   1733 ACTGTGCTGCTCGGGCCCCATTGACTTTTAT  1758
```

RESULT 12
ABK85018/c
ID   ABK85018 standard; DNA; 43545 BP.
XX
AC   ABK85018;
XX
DT   13-AUG-2002 (first entry)
XX
DE   Human cadherin-like asymmetry protein-2 (CLASP-2) genomic DNA.
XX
KW   Human; autoimmune disease; haematopoietic disorder; Digeorge syndrome;
KW   blood protein disorder; agammaglobulinaemia; dysgammaglobulinaemia;
KW   ataxia telangiectasia; common variable immunodeficiency; lymphopenia;
KW   thrombocytopenia; haemoglobinuria; Addison's disease; Grave's disease;
KW   haemolytic anaemia; multiple sclerosis; rheumatoid arthritis; lupus;
KW   endometriosis; autoimmune thyroiditis; anaphylaxis; hypersensitivity;
KW   autoimmune pulmonary inflammation; organ rejection; inflammation;
KW   CLASP; gene; ds.
XX
OS   Homo sapiens.
XX
XX
PN   WO20023111/7-A2.
XX
PD   18-APR-2002.
XX
PF   15-OCT-2001; 2001WO-US32202.
XX
PR   13-OCT-2000; 2000US-0687837.
XX
PA   (ARBO-) ARBOR VITA CORP.
PA   (GARM/) GARMAN J D.
PA   (CAND/) CANDIA A F.
XX
PI   Lu PS;
XX
XX
DR   WPI; 2002-416861/44.
XX
PS   Disclosure; Figure 12B; 245pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) comprising an amino
CC   acid sequence that has 90 % sequence identity to one of the human
CC   cadherin-like asymmetry protein(s) (CLASP)-2 (hCLASP-2A, 2B, 2C, 2E)
CC   sequences (PS). (I) is useful for identifying a compound or agent that
CC   binds CLASP-2 polypeptide. An antibody (III) to (I) is useful for
CC   detecting a CLASP-2 polypeptide in a sample. A pharmaceutical composition
CC   comprising a nucleic acid encoding (I), or (II) is useful for preventing
CC   or treating a CLASP-2 mediated disease e.g. an autoimmune disease, where
CC   the autoimmune disease is caused or exacerbated by increased activity
CC   of TH1 cells. CLASP-2 polynucleotides are useful as probes or primers for
CC   detection or inhibition of CLASP-2 expression (e.g., antisense or
CC   ribozyme-mediated inhibition), for gene knockout, etc. The CLASP-2
CC   polynucleotides can express CLASP-2 polypeptides, produce anti- CLASP-
CC   antibodies or are used as therapeutic polypeptides. The CLASP-2
CC   polynucleotide or fragments can be used in diagnostics (e.g., as probes
CC   for CLASP-2 expression), as a lymphocyte marker and for therapeutic
CC   purposes. CLASP-2 polynucleotide can construct transgenic and knockout
CC   animals, e.g., for screening of CLASP-2 agonists and antagonists. CLASP-2
```

```
CC   polynucleotides can screen for ClASP-2 agonists and antagonists.
CC   ClASP-2 polypeptides or polynucleotides can treat deficiencies or
CC   disorders of the immune system, by activating or inhibiting the
CC   activation, differentiation of immune cells and can treat or detect
CC   deficiencies or disorders of haematopoietic cells. ClASP-2 polypeptides
CC   or polynucleotides can increase differentiation and proliferation of
CC   haematopoietic cells, including the pluripotent stem cells to treat those
CC   disorders associated with a decrease in certain (or many) types of
CC   haematopoietic cells e.g., immunologic deficiency syndromes including
CC   blood protein disorders (e.g. agammaglobulinaemia,
CC   dysgammaglobulinaemia, ataxia telangiectasia, common variable
CC   immunodeficiency, Digeorge syndrome, lymphopenia, thrombocytopenia, or
CC   haemoglobinuria). ClASP-2 polynucleotides or polypeptides can treat or
CC   detect autoimmune diseases, e.g., Addison's disease, haemolytic anaemia,
CC   Grave's disease, multiple sclerosis, rheumatoid arthritis, lupus,
CC   endometriosis, autoimmune thyroiditis, and autoimmune pulmonary
CC   inflammation. ClASP-2 can be used to treat anaphylaxis or
CC   hypersensitivity to an antigenic molecules, organ rejection or graft-
CC   versus-host disease (GVHD) and inflammation. ABK84922-ABK85018
CC   represent cadherin-like asymmetry protein (ClASP) coding sequences
CC   and PCR primers of the invention.
XX
SQ   Sequence 43545 BP; 12383 A; 8378 C; 8983 G; 13801 T; 0 other;
```

```
Query Match             35.6%;  Score 513.4;  DB 24;  Length 43545;
Best Local Similarity   68.0%;
Matches  912;  Conservative    0;  Mismatches  351;  Indels   78;  Gaps   11;
                     68.0%;  Pred. No. 3.2e-126;
```

```
Qy        2 TTTTTTTTTTGAGATGGAGTTTTTTCGCTCTGTTGTTGCCCAGGCTGGAGTGCAATGGCGCAAT 61
            ||||||| |||||||||||||  || ||||||||| ||||||||||||||| |||||||| ||
Db     6705 TTTTTTTTTTTTTGAGATGGAGTTCTT--GCTCTGTTGCCCAGGCTGCAATGGCGGAT 6648

Qy       62 CTCAGCTCACCGCAACCTCCGCCCTCCCGGGTTCAAAGCGACTTCTCCTGCCTCAGCCTCCCG 121
            ||||||||| ||||||||||||||||| |||||||| |||||| | ||||||||||||||||||
Db     6647 CTCGGCCTCACTGCAACCTCCGCCACCCGGGTTCAAGCCGAATTCTCCTGTCTCAGCCTCCCG 6588

Qy      122 AGTAGCTGGGATTACAGGCATGCACCACCACGCTCGGCTAATTTTTGTATTTTTTAGTA 181
            ||| |||||||| |||||| || |||| |||||||||||| ||||||| ||||||||| |
Db     6587 AGTAGCTGGGAATTACAGGCATGCGCCACCACGCCCGGGTTAATTTTTTGTATTTTTTAGTAG 6532

Qy      182 AGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCC 241
            |||||||||||| ||||||||||| |||||||||||||||||||||||| ||||||||||
Db     6531 AGAGACGGGGTTTCTCCATGTTGATCAGGCTGGTCTCGAACTCCTAAATCTCAGGTGATCC 6472

Qy      242 CTCCGTCTCGGCCTCCCAAAGTGCTAGATACA---GGACTGGCCACCATGCCCGGGCTCTG 298
            |||||| |||||||||||||||| || |||||    |||||||||| || |||||||||||
Db     6471 ACCCGCTTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCAGGCCATT 6412

Qy      299 CCTGGCTAATTTTTGTTGGTGGTAGAAACAGGG-------TTTCACTGATTGTGCCCAAGC 347
            |||| |||||  | | || ||| |||  |||||       |||||| |||| |||||||||
Db     6411 ACCCGCTTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATGCCCAGCCATT 6352

Qy      348 -----------------TGGTCTCCTGAGCTCAAGCGAGTCAACCCACGTGCCTCAGC 383
                             |||||||| |||||||||||  | |||| |||||||||||
Db     6351 ATCACCACCATCCTCGTTCGTTGTTCACTCTTGTTCACTGCCACATCCCAACC 6292

Qy      384 CTCCCAAAG-------TGCTGGGATTACAGGCGTGCAGCCGTGCCCGGCCTTTTTTATTTTT 436
            |||||||||       || || |||||| || | |||| ||||||||| |||| | ||||||
Db     6291 GCCTAAGACATAGATAACTAGACACAGAAGATGCTCANATAAATTGCTAAATGAA 6232

Qy      437 ATTTTTTTTAGACACAGGGTGTCCCACTCTTACCCAGGATGGAGTGCAGTGGTGTGATCA 496
            | ||| |||| || |||||||| ||||||||||||||| || | |||| || || |||||
Db     6231 GCCTAAGACATAGATAACTAGACACAGAAGATGCTCAGTGGTGTGATCA 6172

Qy      497 CAGCTCACTGCAGCCTTCAACCTCTGGGATAACAAGCATCTCCTGCCTCAGCCTCCCAAGT 556
            |||||||||||| |||| |||||| |||  ||  ||| | ||||||||||||||||||||||
Db     6171 GGCTTAAATGCTTCTTTTAAAATAAATGACAGGGCAGGGGTAGCATCCTGTCTACCTCAA 6112

Qy      557 AGCTGGGACCAAAGACATGCACCACCACCACTGGCTAATTTTTGTAGTTTTAGTTTATTTTTA 610
            |||||||| | |||||| ||| || || |||||||||||||| | |||||||| |||||| |
Db     6111 TAACTAGGCTTAAGTTTCCTTTTTTCTTTTAAATTTTTATTTTATTTTATTTTA 6052

Qy      611 TAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCAAATCT 670
            | |||||| |||||| |||||||||  ||||||||||||||||||||| ||| ||||| |||
```

```
Db     6051 TTTTATTTTTTGAGACAGAGTCTCGGCTCTGTCGCCCCAGGCTGGAGTGCAGTGGCGCCAATCT 5992
            |||| |  | | ||||||||||||  |||||||||||||||||||||||||||||||| ||||||
Qy      671 TGGCCTCACTGCAACCTCTGCCCTCCGGGTTCAAGTTATTCCTCCGCCCCCAGCCTCCGAG 730
            ||| |||||| |||||||| ||||| |||||||||| ||| |||||||||||||||||||||
Db     5991 CAGCTCACTGCACCTCGCCCTCCGCCTCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAG 5932

Qy      731 TAGCTGGGACTACAGGCGCCCACCACCACGCCTAGCTA--ATTTTTTTGTATTTTTTAGTAGAG 788
            ||||||||||||||||||| ||||||||||| |||| |  |||||||| |||||||||||||||
Db     5931 CAGCTGGGACTACAGGCGCCCACCACGCCCGGGCTAATTTTTTGTATTTTTTAGTAGAG 5872

Qy      789 ATGGGG-TTCACCATGTTGGCCCAGGTTGATCTTGAACTCCTGGGACTTGTGATCTGCCCTGC 847
            |||||| |||||| |||| | ||||| | |||||||||||||||||| | |||| ||||| |
Db     5871 ACGGGGTTTCACCGGTGTTAGCCAGGAGGTTCAATCTCCTGACCTCGTGATCTGCCCAC 5812

Qy      848 CTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTTATTTTTAA 907
            ||||||||||||||||||||||||||||||| |||||||||||||||| || |||| |||| |
Db     5811 CTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGCCCGGCCTAATT---5756

Qy      908 TTTTTGTTTGTTTGAGATGGAGTCTCACTCTGTTACCCAGGCTGGAGTCAATGGCCAAA 967
            |||||| | ||| |||| |||||||||||| | ||||||||||||||| |||| |||||
Db     5755 TTTTAATTTTTTGTAGACGGAGTCTCTCGTCTGTCACCAGGCTGGAGTGCAGTGGCGCATGA 5696

Qy      968 TCTCGGCTCACTGCGAACCTCTGCCTCCCGGGGTTCGGCCGATTCTCCTGCCTCAGCCTCCC 1027
            |||||||||||||||||||||||||| | |||||| || ||||||||||| ||||||||||
Db     5695 TCTCGGCTCACTGCGAACCTCTGCCTCCCCAAGTTCAAGTGATTGTCATGCTTCAGCCTCCT 5636

Qy     1028 AAGCAGCTGGGATTACAGGCGCCCGCACCACCACACCCCGCTAATTTTTGTATTTTTAGTAG 1087
            |||  |||||||| |||||||| | ||||| ||| |||||| |||||||| |||||| ||||
Db     5635 GAGTAGCTGGGACTACAGGCGCGCACCACCACCCGCTAATTTTTGTATTTTTAGTAC 5576

Qy     1088 AGGCGGGGGTTTCACCATATTGTCAGGCTGGTCTGGAACTCCTGACCTCAGGTGACCGAC 1147
            ||| ||||||||||||| | || |||||||||||  |||||| ||||||| |||| |||||
Db     5575 AGATGGGGGTTTCACCACATGTTGGCCAGGCTGGTCTCGAACTCCTGAACTCAAGCAATCCAC 5516

Qy     1148 CTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCTCACCCAGCCGGCTA 1207
            |||||||||| ||||||||||| ||| ||| ||||| ||||||||| |||||  ||||| ||
Db     5515 CTGCCTCACACTCCCAAAGCACTAGGATTACAGGTGTGAGCCACCACGCCCAGCCTCTCA 5456

Qy     1208 ATTTTAGATAAAAAATATGTAGCAAATGGGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCA 1267
            | ||||||| | |||| ||| ||||||||||||| |||||| ||||| |||| |||||||||
Db     5515 CTGCCTCACACTCCCAAAGCACTAGGATTACAGGTGTGAGCCACCACGCCCAGCCTCTCA 5410

Qy     1268 AACTTCTTGGCTTCATGCAATC 1288
            |||| |  ||||||| ||| |||
Db     5409 AACTCCTGGGCTCAAACAATC 5389
```

```
RESULT 13
ABK42270
ID   ABK42270 standard; DNA; 21470 BP.
XX
AC   ABK42270;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Genomic sequence #169 encoding novel human connective tissue polypeptide.
XX
KW   Human; connective tissue related disorder; cancer; gene therapy;
KW   cytostatic; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155343-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01322.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
```

PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 30-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0227182.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 05-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232080.
PR 14-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 26-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 01-NOV-2000; 2000US-0244618.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 08-NOV-2000; 2000US-0249244.
PR 08-NOV-2000; 2000US-0249245.
PR 08-NOV-2000; 2000US-0249264.
PR 08-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 05-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-DEC-2000; 2000US-0256719.
PR 05-DEC-2000; 2000US-0251988.
PR 05-JAN-2001; 2001US-0256978.

XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-565190/63.
XX
XX
PT Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder
PT such as cancer or rheumatoid arthritis.
XX
PS Disclosure; SEQ ID NO 1157; 673pp; English.
XX
CC The present invention relates to the isolation of novel human connective

CC tissue related polypeptides (AAU86435-AAU86223) and the polynucleotide
CC (cDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful
CC in gene therapy. ABKA2102-ABK43116 represent genomic sequences encoding
CC the novel human connective tissue related polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 21470 BP; 5524 A; 4783 C; 5140 G; 6023 T; 0 other;

Query Match 35.6%; Score 513; DB 23; Length 21470;
Best Local Similarity 71.2%; Pred. No. 3.1e-126;
Matches 885; Conservative 0; Mismatches 290; Indels 68; Gaps 13;

Qy   4 TTTTTTTTTTCGAGATGGAGTTTTTCGGCTCTTCGTTGTTGCCCAGGCTGGAGTGGCCAATCT 63
          || ||||||||| ||||||||||| ||||||||||||||||||||||||||||||||||||
Db   2359 TTAATTTTGAGATGCCAGTTTT--GCTCTTGTTCCCTAGGCTGGAGTGCAATGGCACGGATCT 2417

Qy   64 CAGGTCACCGCAACCTCCGGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG 123
          ||| |||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   2418 TGGGTCGTTGCAACCTCTCGCCTCCTGGGTTCAACGCGATTCTCCTGCCTCAGCCTCCCTTAG 2477

Qy   124 TAGCTGGGATTACACAGGCATGTGCCACCACGCCCGGCTAATTTGTATTTTTAGTAGAG 183
          ||||||||||||||| ||||||||||||||||||||||||| ||||||||||||||
Db   2478 TAGCTGGGATTATAGGCACGTGCCACCATGCCTGGCTAATTT---TGTATTTTTAATAG 2534

Qy   184 AGATTGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGTGATCCCT 243
          |||||||||||| || |||||||||||||||||||||||||||| ||||||||||||||
Db   2535 AGACGGGGTTTCGCCATATTGGCCAGGCTGTTCTCGAACTCCTGACCTCAGGTGATCCGC 2594

Qy   244 CCGTCTCTGGGCCCTCCCCCAAAGTGCT-----------AGATACAGGACTGGCCACCATGCCCGCC 294
          ||||||||||||||||||||||||||| || |||||||||||||||||||| ||
Db   2595 CCAACTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGGGCGCCACGGCCGGCCTGGCCTTCCA 2654

Qy   295 TCTGCCTCGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGT--------------- 339
          ||| | |||||||||||| |||||||||||||||||||||||
Db   2655 TTATAGAATTAACCATAATTATTTAATCACAATAGCAGAAATACTAGATTGAAATCGAAT 2714

Qy   340 ----GCCCAAAGCTGGTCTCCTGAGCTCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGC 395
              ||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   2655 TTATAGAATTAACCATAATTATTTAATCACAATAGCAGAAATACTAGATTGAAATCGAAT 2714

Qy   396 TGGGATT------ACAGGCGTGCAGCCGCGTGCCCGGCCTTTTTAATTTTATTTTTTTTTTT 450
          ||||||| ||| ||||||||| ||||||||||||||||||| |||||||||||||
Db   2715 TCTCCTTCCCTCTTTAACTGCTGGGCCTTACTACCCTCCAGATGCAGTCAGTCTACACAATGAA 2774

Qy   451 ACAGGTGTCCCACTCTTACCCAGGATGAAGTGCAGTGGTGTGATCACACTGCTCACTGCAGC 510
          || |||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db   2835 GAGAGTCTCACTCTGTTGCCCAGGCTGGAGTGCACAATCTTGGCTCACTGCAAC 2894

Qy   511 CTTCAACTCCTGAGATCAAGC-ATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAA 569
          ||||||||||||||||||||| |||||||||||||||||||||||| ||||||||||||
Db   2895 CTCTGCCTCCCGGATTCAAGCGAATTCTCCTCACCTCCCAACTAGCTGGGA-TTAC 2953

Qy   570 GACATGCACCACTACACCTGCCTAATTTTTTTATTTTTTAGTAGAGACAGAGT 629
          |||||||||||||||||||| ||||||| ||||||||||||||||||||||||
Db   2954 GGTGTGTGCCACCACGGCCTGGCCAGATTTT-----------TTTTTTGAGACGGAGT 2998

Qy   630 CTCAACTCCTGTCACCCAGGCTGGAGTGCAGTGGCAAATCTTGGCTCACTGCAAC 689
          || |||||||||||||||||||||||||||||||||||||||||| ||||||
Db   2999 CTT-GCTCTGTCACCAGGCTGGTGGTGGCCATCTAGGGTCACTGAACCTCCG 3057

Qy   690 CCTCCGGGTTCAAGTAATTCCTGCCCCAGCCTCTCGAGTAGCTGGGACTACAGGCGC 749
          ||||| ||||||||||| || |||||||||||||||||||||||||||||||||||
Db   3058 CCTCCCAGGTTCATGCAATTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCAT 3117

Qy   750 CCACCACGCCTAGCTAAT-TTTTTTTGTATTTTTTAGTAGAGATGGGGTT---CACCATGTTGG 804
          |||||||||| ||||||| |||||||||||||||||||||||||||||| |||||||||||
Db   3118 GTATCACTACGCGCCCAGGTAAATTTTTTGTATTTTTTAGTAGAGACAGGGTTTCACCATGTTTGG 3177

Qy   805 TCGCGCAGGGTTGATCTTGATCTCTGGACCTTGTGATCTCGGCGGCCTCCCAAAGTG 864
          | |||| ||||||||||||||||||| ||||||||||||||||||||| |||||||||
Db   3178 GGGTCAGGCTGGTCTTGAACTCCTGACCTCGTGATCCCGCCTCAGCCTCCCAAAGTG 3237

Qy   865 CTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTATTTTTAATTTTTTGTTTTGTTTGAGA 924
          |||||||||||||||||||||||||||||||| ||||||||||||||||| |||| |||||||
Db   3238 CTGGGATTACGGGCGTGAGCCACCGCCGCCGGCCT-----AATTTTTTTTTGTTTTTTTGAGA 3290

Qy   925 TGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAATGGCAAAATCTCGGCTCACTGCAAC 984
          |||||||||||| || ||| |||||||||||| |||| ||||| |||| ||||||||||||
Db   3291 CAGGAGTCTCACTCCTTCATCCAGGCTAGAGTGCAGTGGTGTGATCTCGGCTCACTGCAAC 3350

Qy   985 CTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGCAGCTGGGATTACG 1044
          |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db   3351 CTCTGCCTTCCCGGGTTCAAGCAATTCTCCTGCCTCCAGTCCCCGGTTCAGCTGGGATTACA 3410

Qy   1045 GGCACCTGCCACCACACGCCGCTAATT-TTTGTATTTTTCATTAGAGGCGGGGTTTCACCA 1103
          ||||||||||||||||||||| ||||| ||||||||||||||| ||||||||||||||||
Db   3411 GGCGCGTGCCACCATGCCTGGCTAATTTTTTGTAATTTTTAGTAGAGACGGGGTTTCACCA 3470

Qy   1104 TATTTGTCAGGCTGGTCTCGAAACTCCTGACCTCAGGTGATCACCCGCTGCTCAGCCTTCCA 1163
          |||||||||||||||||||||| ||||||||||||||| ||||| |||||||||||||||||
Db   3471 CATTGGCCAGGCTGGTCTCAAAACTCCTGACCTCAAGTGATCCACCCGCCTCGGCCTCCCA 3530

Qy   1164 AAGTGCTGGGATTACAGGCGTGAGCCACCGCACCTCACCCAGCCGGGCT 1206
          ||||||||||||||||||||| |||||||||||||||| ||| | |
Db   3531 AAGTGCTGGGATTACAGGCGTGGGCCCACTGCACCTGGCCTGTT 3573

RESULT 14

AAK71768
ID   AAK71768 standard; DNA; 5262 BP.
XX
AC
XX
DT   06-NOV-2001 (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26580.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   14-JUL-2000; 2000US-0217496.
PR   26-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   14-AUG-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.

```
PR   14-AUG-2000;   2000US-0225267.
PR   14-AUG-2000;   2000US-0225268.
PR   14-AUG-2000;   2000US-0225270.
PR   14-AUG-2000;   2000US-0225447.
PR   14-AUG-2000;   2000US-0225757.
PR   14-AUG-2000;   2000US-0225758.
PR   14-AUG-2000;   2000US-0225759.
PR   18-AUG-2000;   2000US-0226279.
PR   22-AUG-2000;   2000US-0226681.
PR   22-AUG-2000;   2000US-0226868.
PR   22-AUG-2000;   2000US-0227182.
PR   23-AUG-2000;   2000US-0227009.
PR   30-AUG-2000;   2000US-0228924.
PR   01-SEP-2000;   2000US-0229287.
PR   01-SEP-2000;   2000US-0229343.
PR   01-SEP-2000;   2000US-0229344.
PR   01-SEP-2000;   2000US-0229345.
PR   05-SEP-2000;   2000US-0229509.
PR   05-SEP-2000;   2000US-0229513.
PR   06-SEP-2000;   2000US-0230437.
PR   06-SEP-2000;   2000US-0230438.
PR   08-SEP-2000;   2000US-0231242.
PR   08-SEP-2000;   2000US-0231243.
PR   08-SEP-2000;   2000US-0231244.
PR   08-SEP-2000;   2000US-0231413.
PR   08-SEP-2000;   2000US-0231414.
PR   08-SEP-2000;   2000US-0232080.
PR   08-SEP-2000;   2000US-0232081.
PR   12-SEP-2000;   2000US-0231968.
PR   14-SEP-2000;   2000US-0232397.
PR   14-SEP-2000;   2000US-0232398.
PR   14-SEP-2000;   2000US-0232399.
PR   14-SEP-2000;   2000US-0232400.
PR   14-SEP-2000;   2000US-0232401.
PR   14-SEP-2000;   2000US-0233063.
PR   14-SEP-2000;   2000US-0233064.
PR   14-SEP-2000;   2000US-0233065.
PR   21-SEP-2000;   2000US-0234223.
PR   21-SEP-2000;   2000US-0234274.
PR   25-SEP-2000;   2000US-0234997.
PR   25-SEP-2000;   2000US-0234998.
PR   26-SEP-2000;   2000US-0235484.
PR   27-SEP-2000;   2000US-0235834.
PR   27-SEP-2000;   2000US-0235836.
PR   29-SEP-2000;   2000US-0236327.
PR   29-SEP-2000;   2000US-0236367.
PR   29-SEP-2000;   2000US-0236368.
PR   29-SEP-2000;   2000US-0236369.
PR   29-SEP-2000;   2000US-0236370.
PR   02-OCT-2000;   2000US-0236802.
PR   02-OCT-2000;   2000US-0237037.
PR   02-OCT-2000;   2000US-0237038.
PR   02-OCT-2000;   2000US-0237039.
PR   02-OCT-2000;   2000US-0237040.
PR   13-OCT-2000;   2000US-0239935.
PR   13-OCT-2000;   2000US-0239937.
PR   20-OCT-2000;   2000US-0240960.
PR   20-OCT-2000;   2000US-0241221.
PR   20-OCT-2000;   2000US-0241785.
PR   20-OCT-2000;   2000US-0241786.
PR   20-OCT-2000;   2000US-0241787.
PR   20-OCT-2000;   2000US-0241808.
PR   20-OCT-2000;   2000US-0241809.
PR   20-OCT-2000;   2000US-0241826.
PR   01-NOV-2000;   2000US-0244617.
PR   08-NOV-2000;   2000US-0246474.
PR   08-NOV-2000;   2000US-0246475.
PR   08-NOV-2000;   2000US-0246476.
PR   08-NOV-2000;   2000US-0246477.
PR   08-NOV-2000;   2000US-0246478.
PR   08-NOV-2000;   2000US-0246523.
PR   08-NOV-2000;   2000US-0246524.
PR   08-NOV-2000;   2000US-0246525.
PR   08-NOV-2000;   2000US-0246526.
PR   08-NOV-2000;   2000US-0246527.
PR   08-NOV-2000;   2000US-0246528.
PR   08-NOV-2000;   2000US-0246532.
PR   08-NOV-2000;   2000US-0246609.
PR   08-NOV-2000;   2000US-0246610.
PR   08-NOV-2000;   2000US-0246611.
PR   08-NOV-2000;   2000US-0246613.
PR   17-NOV-2000;   2000US-0249207.
PR   17-NOV-2000;   2000US-0249208.
PR   17-NOV-2000;   2000US-0249209.
PR   17-NOV-2000;   2000US-0249210.
PR   17-NOV-2000;   2000US-0249211.
PR   17-NOV-2000;   2000US-0249212.
PR   17-NOV-2000;   2000US-0249213.
PR   17-NOV-2000;   2000US-0249214.
PR   17-NOV-2000;   2000US-0249215.
PR   17-NOV-2000;   2000US-0249216.
PR   17-NOV-2000;   2000US-0249217.
PR   17-NOV-2000;   2000US-0249218.
PR   17-NOV-2000;   2000US-0249244.
PR   17-NOV-2000;   2000US-0249245.
PR   17-NOV-2000;   2000US-0249264.
PR   17-NOV-2000;   2000US-0249265.
PR   17-NOV-2000;   2000US-0249297.
PR   17-NOV-2000;   2000US-0249299.
PR   17-NOV-2000;   2000US-0249300.
PR   01-DEC-2000;   2000US-0250160.
PR   01-DEC-2000;   2000US-0250391.
PR   05-DEC-2000;   2000US-0251030.
PR   05-DEC-2000;   2000US-0251988.
PR   06-DEC-2000;   2000US-0256719.
PR   06-DEC-2000;   2000US-0251479.
PR   08-DEC-2000;   2000US-0251856.
PR   08-DEC-2000;   2000US-0251868.
PR   08-DEC-2000;   2000US-0251869.
PR   08-DEC-2000;   2000US-0251989.
PR   08-DEC-2000;   2000US-0251990.
PR   11-DEC-2000;   2000US-0254097.
PR   05-JAN-2001;   2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA, Barash SC,   Ruben SM;
XX
DR   WPI: 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
     useful for preventing, diagnosing and/or treating cancers and
     metastasis -
PS   Disclosure; SEQ ID NO 26580; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK54942
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54951 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 5262 BP; 1099 A; 1347 C; 1247 G; 1569 T; 0 other;
```

```
Query Match    35.5%;  Score 511.6;  DB 22;  Length 5262;
Best Local Similarity   70.3%;  Pred. No. 4.2e-126;
Matches  922;  Conservative  0;  Mismatches  299;  Indels  91;  Gaps  14;

Qy       1 TTTTTTTTTTTTTGAGATGGAGTTTTTTGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA     60
           ||||  |||||| ||||||||||||| ||||||||||||||||||||||||||| | ||||||
Db    1381 TTTTATGTTTTGAGACAGAG-TCTCTCTCTTGTTGCCCAGGCTGGAATGCAATGGCGTGA   1439

Qy      61 TCTCAGCTCACCGCAACCTCCGCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC    120
           ||||| |||||||| |||||| ||||||| ||||| ||||||| | ||||||||||||| |
Db    1440 TCTCGGCTCACTGCAGCCTCCGTCCCAGGTTCAAGCGATTCTCATGCCTCAGCCTCCC   1499

Qy     121 CAGTAGCTGGGATTACAGGCACATGTGACCACCACGCCCGGCTAATTTTGTATTTTTTTAG    180
           |||||||| |||||||||||| |||||||| ||| |||||||||||||||||||||| |||
Db    1500 CAGTAGCTAGGGATTACAGGCGCACGTGCCACCATGCCCAGCTAATTTT---TGTATTTTTG   1556

Qy     181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC    240
           ||||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    1557 TAGAGACGGGGTTTCACCATGTTGACCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATC   1616

Qy     241 CCTCCGGTCTCGGCCTCCCAAAAGTGCTAGGATACAGGACTGGCCACCACGCCCGGCTCTGCC    300
           |||||||||||| |||||||||||||||||||| | || ||||||||||||||||||||||
Db    1617 TGCCCGGCTTGGGCCTCCCAAAG-TGCTGGGATTACAGGCGTGAACTACCGGCCCTGGC   1675

Qy     301 TGGCTAATTTTGTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC    360
           |||||||||||||||||| |||||||||||||| |||||||||| |||||||||||| ||
Db    1676 CTGCATGTTTCTTTAATAGCATCCTGTCTGTTTTGTTTAATTTCTGTATGGGACTTCTTT-   1728

Qy     361 TCAAGCAGTCCACTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC    420
           |||||||| ||||| ||||||||||||||||||||||||||| | |||||| |
Db    1729 GACAGTGGCATGATCGCACACTCACTGCACCCAATCTCCCAGGTTCAAGTGATTCTCCT   1787

Qy     421 CTGGCCTTTTTTATTTTATTTTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG    480
           ||||||||||||||||| |||||||||||||||||||||||||| |||||||||||||||
Db    1788 GCCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCACCACCACTACACCTGGCTAATTTTT   1847

Qy     481 TGCAGTGGTGTGATCACAGCTCACTGCACCTTCAACTCCTGAGATCAAG-CATCCTCCT    539
           |||||||||||||||||||||||||| || |||||||| ||||| |||| ||||||||||
Db    1848 GCCTCAGCCTCCCAGTAGCTGGGATTACAGGCATGGCCACCACACCCAGCTAATTTTT   1907

Qy     540 GCCTCAGCCTCCCAAGTAGCTGGGACTAGCTGGGGACTACAGGCGCCCAGCCACCAT    599
           |||||||||||||||||||||||| || ||||| |||||||||||||||||
Db    1908 -TGTTTGTTTGTTTGTTTGTTTTTGAGGAGTTTC-ACTCGTCACCCAGGCTGGAGTGCAG   1965

Qy     600 ATTTTTTATTTTTTATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGCTGGAGTGCAG    659
           |||||||||||||||||||||||||||||||| |||||||| |||||||||||||||||
Db    1966 TCGTGCCATGTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGCTTCTCCTGTCTC   2025

Qy     660 TGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCGTCGCCC    719
           ||||||||||||||||||||||||||| |||||||||||||||||| | || | |||||
Db    2026 AGCCTCCTGAGTAGCTGGGACTACAGGCGCCACCACGCCTAGCTAAT-TTTTGTATTTT   2085

Qy     720 AGCCTCCTGAGTAGCTGGGACTACAGGCGCCACCACGCCTAGCTAAT-TTTTGTATTTTATT    778
           ||||||||||||||||||||||||||||||||||||||||||||||| |||||| ||||
Db    2086 TTTAGTATAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGG   2145

Qy     779 TTTTAGTAGAGATGGGG-TTCACCAGGTTGGTCCCAGGTTGATCTTGGGACCT-TG    835
           ||||| ||||| ||||||||||||||||||||||||||||||| |||||| ||
Db    2146 TGATCCGGCCGCCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACT   2205

Qy     836 TGATCTGCCGCCTGCCTCGGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCGCACC    895
           |||||| ||||||||||| |||||||||||| ||||||||||||||||||||| ||||||
Db    2206 GCCTGTT-----TGTTTCTTTTTTTTGTTTGTTTTGAAATGGAAATCTCACTC   2260

Qy     896 GCTTATTTTTAATTTTGTTGTTGTTTTGAAATGGAAATCTCACTC---TGTTACCCAGGCTGG    952
           |||||||||| ||||| |||||||||| ||||||||||||||||    ||||||||||||||
Db    2261 AGTGCAGTGGCAATCTCGGCTCACTGCAACCTCCACTT---TGGGTTCAAGCGATTCTC   2318

Qy     953 AGTGCAAATGCCAAATCTCGGCTCACTGCAACCGATTCTC   1012
           ||||||||||||||||||||||||||||||||||  |||||
Db    2319 CTGCCTCAGCCT-CGAAGTAGCTGGGATTACAGGCATGTGCCACCACACCTGGTTAAATT   2377

Qy    1013 CTGTCTCAGCCTCCCAAGAGCAGTCGGGATTACGGGGGCACCCGGCTAATTT   1072
           ||||||||||||||| | |||   || |||||| |    ||||  |
Db    2378 TTGTATTTTTCATTAGAGACGGGGTTTCACCATATATTTGTCAGGCGTCTCAAAACTCCTGA   2437
```

```
PR   22-AUG-2000;   2000US-0226868.
PR   22-AUG-2000;   2000US-0227182.
PR   23-AUG-2000;   2000US-0227009.
PR   30-AUG-2000;   2000US-0228924.
PR   01-SEP-2000;   2000US-0229287.
PR   01-SEP-2000;   2000US-0229343.
PR   01-SEP-2000;   2000US-0229344.
PR   01-SEP-2000;   2000US-0229345.
PR   05-SEP-2000;   2000US-0229509.
PR   05-SEP-2000;   2000US-0229513.
PR   06-SEP-2000;   2000US-0230437.
PR   06-SEP-2000;   2000US-0230438.
PR   08-SEP-2000;   2000US-0231242.
PR   08-SEP-2000;   2000US-0231243.
PR   08-SEP-2000;   2000US-0231244.
PR   08-SEP-2000;   2000US-0231413.
PR   08-SEP-2000;   2000US-0231414.
PR   08-SEP-2000;   2000US-0232080.
PR   08-SEP-2000;   2000US-0232081.
PR   12-SEP-2000;   2000US-0231968.
PR   14-SEP-2000;   2000US-0232397.
PR   14-SEP-2000;   2000US-0232398.
PR   14-SEP-2000;   2000US-0232399.
PR   14-SEP-2000;   2000US-0232400.
PR   14-SEP-2000;   2000US-0232401.
PR   14-SEP-2000;   2000US-0233063.
PR   14-SEP-2000;   2000US-0233064.
PR   14-SEP-2000;   2000US-0233065.
PR   21-SEP-2000;   2000US-0234223.
PR   21-SEP-2000;   2000US-0234274.
PR   25-SEP-2000;   2000US-0234997.
PR   25-SEP-2000;   2000US-0234998.
PR   26-SEP-2000;   2000US-0235484.
PR   27-SEP-2000;   2000US-0235834.
PR   27-SEP-2000;   2000US-0235836.
PR   29-SEP-2000;   2000US-0236327.
PR   29-SEP-2000;   2000US-0236367.
PR   29-SEP-2000;   2000US-0236368.
PR   29-SEP-2000;   2000US-0236369.
PR   29-SEP-2000;   2000US-0236370.
PR   02-OCT-2000;   2000US-0236802.
PR   02-OCT-2000;   2000US-0237037.
PR   02-OCT-2000;   2000US-0237038.
PR   02-OCT-2000;   2000US-0237039.
PR   02-OCT-2000;   2000US-0237040.
PR   13-OCT-2000;   2000US-0239935.
PR   13-OCT-2000;   2000US-0239937.
PR   20-OCT-2000;   2000US-0240960.
PR   20-OCT-2000;   2000US-0241221.
PR   20-OCT-2000;   2000US-0241785.
PR   20-OCT-2000;   2000US-0241786.
PR   20-OCT-2000;   2000US-0241787.
PR   20-OCT-2000;   2000US-0241808.
PR   20-OCT-2000;   2000US-0241809.
PR   20-OCT-2000;   2000US-0241826.
PR   01-NOV-2000;   2000US-0244617.
PR   08-NOV-2000;   2000US-0246474.
PR   08-NOV-2000;   2000US-0246475.
PR   08-NOV-2000;   2000US-0246476.
PR   08-NOV-2000;   2000US-0246477.
PR   08-NOV-2000;   2000US-0246478.
PR   08-NOV-2000;   2000US-0246523.
PR   08-NOV-2000;   2000US-0246524.
PR   08-NOV-2000;   2000US-0246525.
PR   08-NOV-2000;   2000US-0246526.
PR   08-NOV-2000;   2000US-0246527.
PR   08-NOV-2000;   2000US-0246528.
PR   08-NOV-2000;   2000US-0246532.
PR   08-NOV-2000;   2000US-0246609.
PR   08-NOV-2000;   2000US-0246610.
PR   08-NOV-2000;   2000US-0246611.
PR   08-NOV-2000;   2000US-0246613.
PR   17-NOV-2000;   2000US-0249207.
```

```
PR   17-NOV-2000;   2000US-0249208.
PR   17-NOV-2000;   2000US-0249209.
PR   17-NOV-2000;   2000US-0249210.
PR   17-NOV-2000;   2000US-0249211.
PR   17-NOV-2000;   2000US-0249212.
PR   17-NOV-2000;   2000US-0249213.
PR   17-NOV-2000;   2000US-0249214.
PR   17-NOV-2000;   2000US-0249215.
PR   17-NOV-2000;   2000US-0249216.
PR   17-NOV-2000;   2000US-0249217.
PR   17-NOV-2000;   2000US-0249218.
PR   17-NOV-2000;   2000US-0249244.
PR   17-NOV-2000;   2000US-0249245.
PR   17-NOV-2000;   2000US-0249264.
PR   17-NOV-2000;   2000US-0249265.
PR   17-NOV-2000;   2000US-0249297.
PR   17-NOV-2000;   2000US-0249299.
PR   17-NOV-2000;   2000US-0249300.
PR   01-DEC-2000;   2000US-0250160.
PR   01-DEC-2000;   2000US-0250391.
PR   05-DEC-2000;   2000US-0251030.
PR   05-DEC-2000;   2000US-0251988.
PR   05-DEC-2000;   2000US-0251989.
PR   06-DEC-2000;   2000US-0251479.
PR   06-DEC-2000;   2000US-0255719.
PR   08-DEC-2000;   2000US-0251856.
PR   08-DEC-2000;   2000US-0251868.
PR   08-DEC-2000;   2000US-0251869.
PR   08-DEC-2000;   2000US-0251989.
PR   08-DEC-2000;   2000US-0251990.
PR   11-DEC-2000;   2000US-0254097.
PR   05-JAN-2001;   2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,   Barash SC,   Ruben SM;
XX
DR   WPI;  2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis.
XX
PS   Disclosure;  SEQ ID NO 26581;  307ipp + Sequence Listing;  English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 5262 BP;  1100 A;  1346 C;  1571 G;  1571 T;  0 other;

Query Match             35.5%;  Score 511.6;  DB 22;  Length 5262;
Best Local Similarity   70.3%;
Matches  922;  Conservative   0;  Mismatches 299;  Indels   91;  Gaps  14;

Qy     1 TTTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
         |||||||||||| |||||||||||||||||||| |||||||||||||||||||| |||||| |
Db  1381 TTTTATGTTTTGAGACAGAG-TCTCTCTTGTTGTTGCCCAGGCTGGAATGCAAATGGCGTGA 1439

Qy    61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCC 120
```

```
Db 2438 CCTCAAGTGATCCGACGCGCCGCCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACT 2497
        ||||  |||| | |||         ||  | || ||| ||||||||||||||||||||||||||||
Qy 1193 TCACCCAGCGCCGGCTAATTTTAGATAAAAAATATGTRAGCAAATGGGGGGATTACTTTGCTATGTTG 1252

Db 2498 GCTCCTTGGCCCACCTGGCTAATTTTTGTATTTTTAGTAGAGATGGAGTTTCACCATGTTG 2557
        ||| ||  | ||        |||||||||||||||||||||||||||||||||||||||||
Qy 1253 CCCAGGCTGGTCTCAAACTTCTGGCTTCATCGCAATCCTTCCAAATGAGCCAC 1304

Db 2558 GCCAGGCTGGTCTCGAAGTCCTGACCTCAGGTTATCCACCTGCCTCAGCCCTC 2609
```

Search completed: February 1, 2003, 21:37:50
Job time : 605 secs

```
Db 1440 TCTCGGCTCACTGCAGCCTCCAACTGCCCAGGTTCAAGCGATTCTCATGCCTCAGCCTCCC 1499
        |||||  |||  |||        |||    || |||  |||||  ||| ||||||||||||||
Qy  121 CAGTAGCTGGGATTACAGGCATGTGCACCACCAGCGTCGGCTAATTTTTGTATTTTTTTAG  180

Db 1500 CAGTAGCTAGGATTACAGGCACGTGCCACCATGCCCAGCTTATTTT---TGTATTTTTGG 1556
        ||||||||| |||||||||| | |||||||| ||| |||  ||||||||  |||||||||
Qy  181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC  240

Db 1557 TAGAGACGGGGTTTCACCATGTTGACCAGGCTGGTCTCGAAATCCTGACCTCAAGTGATC 1616
        ||||| |||||||||||||||||||||||||||||||||   ||| ||||||||||||||
Qy  361 TCAAGCAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC  420

Db 1676 CTGACATGTTTCTTTAATAGCATCCTGTTGTTTAATTTCTGTATGGACTTTCTTT----- 1728
        || |||| ||||||||||||||||||||||||||||||||||||| |||||||||
Qy  301 TGGCTAATTTTTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC  360

Db 1617 TGCCCGGCCTTGGCCTCTCCCAAAG-TGCTGGGATTACAGGCGTGAACTACCGCGCCTGGC 1675
        | |||| |||||| ||||||||||| |||||||||||||||||||| |||||| ||||||
Qy  241 CCTCCGTCTCCGGCCTCCCAAAGTGCTAGATACAGGACTGGGCCACCATGCCCGGCTCTGCC  300

Db 1788 GACAGTGGCATGATGCAGCGTCACTGCAACCCCAAATCTCCCAGGTTCAAGTGATTCTCCT 1847
        | |||||| ||| |||| |||||||||| | ||||||||||||||||||||| |||||||||
Qy  481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAG-CATCCTCCT  539

Db 1729 -TGTTTGTTGTTTGTTGTTTGTATGGGAGTTTC-ACTCTGTCACCCAGGCTGGAGTGCAG 1787
        |||||||| |||||| |||||||| || |||||||||||||| ||||||||||||||||
Qy  421 CTGGCCTTTTTTATTTTATTTTTTTTAATTTTTTTAAGACACAGGGTCTCCCAACTCTTACCCAGGATGAAG  480

Db 1848 GCCTCAGCCTCCCAGTAGCTGGGATTACAGGCATGCGACACACCACCACCAGCTAATTTT 1907
        |||||||||||||| |||||||||||||||||| || | ||||||| |||||||||||||
Qy  540 GCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACCACTACACCTGGCTAATTTTT  599

Db 1966 TCGTGCTCCATGTTGGGCTCACTGCAACCTCCGCCTCCCAGGTTCAAGTGCTTCTCCTGTCTC 2025
        |  | |||| ||||||| ||||||||| |||| |||||| |||||||||||||||||||| |||
Qy  600 ATTTTTATTTTTATTTTTTTGAGACAGAGTCTCACTCTGTTGTCCCAGGCTGGAGTGCAG  659

Db 2026 AGCCTCCTGAGTACCTGGGATTATAGGCATGCCACCACATGTCCAGCTAATTTTTGTATT 2085
        ||||||||||| | ||||||||| ||||| |||||||| || ||||||||| ||||||||
Qy  660 TGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCCC  719

Db 2086 TTTAGTAGTAGACGGGGTTTCACCAGTGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGG 2145
        |||||||| |||| |||||||||||| ||||  |||||||||||||| ||| |||||| ||
Qy  720 AGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCCAGCTAAT-TTTTTGTATT  778

Db 2146 TGATCCGGCCTGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCTCC 2205
        |||| |||||||||| |||||||||||||||||||||||||||||||||||||||| |||
Qy  779 TTTAGTAGAGATGGGG--TTCACCAGTGTTGCCCCAGGTTGATCTCGATCTCTGGGACCT--TG  835

Db 2206 GCCCTGTT-----TGTTTCTTTTTTTGAGAAGGAGTCTCACTC 2260
        |||  ||||     |||| || ||||| | |||  || || |||
Qy  836 TGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCG  895

Db 2261 AGTGCAAGTGGAGGGATCTCGGCTCACTGCAACCTCCACTT--TGGGTTCAAGCGATTCTC 2318
        |||||||| || |||  |||||||||||||||||||||||||  | |||||||||||||||
Qy  896 GCCTTATTTTTTAAATTTTTGTTTTGTTTTGAAATGGAAATCTCACTC---TGTTACCCAGGCTGG  952

Db 2319 CTGCCTCAGCCT-CGAAGTAGCTGGGATTACAGGCATGTGCCACCACACCCGGCTAAATTT 2377
        |||||||||||| | |||||||||||||||| |||| ||||||||||| ||||||||| |||
Qy  953 AGTGCAAATGCCAAATCTCGGGCTCACTGCAACCTCCACTCTGCCTCCCGGGGCTCAAGCGATTCTC 1012

Db 2378 TTGTATTTTAGTAGAGACGGGGTATCGCAATGTTGGCCAGGCTGCTCTCAAACTCCTGA 2437
        ||||||||||||||||| |||||| | |||||||||||||||| ||||||||||||||||
Qy 1013 CTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGCACTGCCACCACGCCCGGCTAATTT 1072

Qy 1073 TTGTATTTTCATTAGAGAGCGGGGGTTTCACCATATTGTCAGGCTGGTCTCAAACTCCTGA 1132
        |||||||| |||||||||  ||| |||||||||||||||| |||||||||||||||||||||
Db 2379 GCCAGGCTGGTCTCGAAGTCCTGACCTCAGGTTATCCACCTGCCTCAGCCCTC 2437

Qy 1133 CCTCAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACC 1192
```

OM nucleic - nucleic search, using sw model

Run on: February 1, 2003, 20:06:24 ; Search time 3874 Seconds
                 Copyright (c) 1993 - 2003 Compugen Ltd.
                 GenCore version 5.1.3
                                          (without alignments)
                                          10832.795 Million cell updates/sec

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 tttttttttttgagatggag...........ttaaacaaagcttagagaca 1442
Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 1455102878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
                 1: gb_ba:*
                 2: gb_htg:*
                 3: gb_in:*
                 4: gb_om:*
                 5: gb_ov:*
                 6: gb_pat:*
                 7: gb_ph:*
                 8: gb_pl:*
                 9: gb_pr:*
                 10: gb_ro:*
                 11: gb_sts:*
                 12: gb_sy:*
                 13: gb_un:*
                 14: gb_vi:*
                 15: em_ba:*
                 16: em_fun:*
                 17: em_hum:*
                 18: em_in:*
                 19: em_mu:*
                 20: em_om:*
                 21: em_or:*
                 22: em_ov:*
                 23: em_pat:*
                 24: em_ph:*
                 25: em_pl:*
                 26: em_ro:*
                 27: em_sts:*
                 28: em_un:*
                 29: em_vi:*
                 30: em_htg_hum:*
                 31: em_htg_inv:*
                 32: em_htg_other:*
                 33: em_htg_mus:*
                 34: em_htg_pln:*
                 35: em_htg_rod:*
                 36: em_htg_mam:*
                 37: em_htg_vrt:*
                 38: em_sy:*
                 39: em_htgo_hum:*
                 40: em_htgo_mus:*
                 41: em_htgo_other:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1442	100.0	1442	6	AR051550	AR051550 Sequence
2	1442	100.0	1442	6	AR072690	AR072690 Sequence
3	1442	100.0	1442	6	AR073235	AR073235 Sequence
4	1442	100.0	1442	9	AF010144	AF010144 Homo sapi
c 5	1301.4	90.2	124001	9	AL031295	AL031295 Human DNA
6	1080.2	74.9	1381	6	HS886K2	HS886K2 Sequence
7	1080.2	74.9	1381	6	AR073164	AR073164 Sequence
8	1080.2	74.9	1381	6	AR072619	AR072619 Sequence
9	1080.2	74.9	1381	6	AR051479	AR051479 Sequence
10	592.2	41.1	160714	2	AC087434	AC087434 Pan trogl
11	590.4	40.9	99370	6	AC005057	AC005057 Sequence
12	587.8	40.8	65608	6	AX330738	AX330738 Sequence
13	587.8	40.8	65608	6	AX332242	AX332242 Sequence
14	587.8	40.8	65608	6	AX335496	AX335496 Sequence
15	587.8	40.8	67046	9	U62293	U62293 Human LIM-k
16	587.8	40.8	148525	2	AL137073	AL137073 Human DNA
c 17	578.4	40.1	121272	2	AC090543	AC090543 Homo sapi
18	567.6	39.4	170630	2	AC016905	AC016905 Homo sapi
19	567.6	39.4	201300	9	AL137073	AL137073 Human DNA
20	562.4	39.0	209856	2	AC090543	AC090543 Homo sapi
21	561.4	38.9	173588	2	AC127468	AC127468 Papio cyn
22	559.2	38.8	202235	2	AC117378	AC117378 Homo sapi
23	559.2	38.8	191169	2	AC032038	AC032038 Homo sapi
24	558.2	38.7	202289	2	AC026687	AC026687 Homo sapi
25	556.8	38.6	137737	2	AC010378	AC010378 Homo sapi
26	556.8	38.6	192096	9	AL590133	AL590133 Human DNA
c 27	555	38.5	91927	9	AC004771	AC004771 Homo sapi
28	554.6	38.5	42665	9	AC011559	AC011559 Homo sapi
29	552	38.3	172262	9	AC123979	AC123979 Papio cyn
30	551.8	38.3	211063	2	AC069254	AC069254 Homo sapi
31	551.8	38.2	155949	9	AC091022	AC091022 Homo sapi
32	550.8	38.2	158863	2	AC069249	AC069249 Homo sapi
33	550.8	38.2	94121	2	HS7115309	HS7115309 Human DNA
34	550.8	38.2	193897	2	AL109806	AL109806 Human DNA
35	550.4	38.2	202289	2	AC069486	AC069486 Homo sapi
36	550.2	38.1	175186	2	AC090681	AC090681 Homo sapi
37	548.2	38.0	194578	2	AC026125	AC026125 Homo sapi
38	548.2	38.0	207607	2	AC125611	AC125611 Homo sapi
39	548.2	37.9	180948	2	AC010173	AC010173 Homo sapi
40	548.2	37.9	122146	9	AC114399	AC114399 Homo sapi
41	543.8	37.7	62485	9	AC011736	AC011736 Homo sapi
42	541.2	37.5	171056	2	AL590093	AL590093 Human DNA
43	540.2	37.3	129043	9	AC026318	AC026318 Homo sapi
44	538.2	37.3	191946	9	AL132712	AL132712 Human chr
45	537.2	37.3	110129	9	AL132780	AL132780 Human chr

ALIGNMENTS

RESULT 1
AR051550
LOCUS   AR051550   1442 bp   DNA   linear   PAT 29-SEP-1999
DEFINITION Sequence 120 from patent US 5830670.
ACCESSION   AR051550
VERSION   AR051550.1 GI:5974914
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
            Unclassified.
REFERENCE   1 (bases 1 to 1442)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE   Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5830670-A 120 03-NOV-1998;

```
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="unknown"
ORIGIN
BASE COUNT    314 a    402 c    311 g    415 t

Query Match    100.0%;  Score 1442;  DB 6;  Length 1442;
Best Local Similarity  100.0%;  Pred. No. 0;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 1442;  Conservative 0;

Qy      1 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTTGAGATGGAGTTTTGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCCAA 60

Qy     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy    121 CAGTAGCTGGGATTACAGGCGCATGCCACCACGCGCCGGCTAATTTTTGTATTTTTAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCGCATGCCACCACGCGCCGGCTAATTTTTGTATTTTTAG 180

Qy    181 TAGAGATGGAGTTTCTCCAAGTTGGTCAGCGCTGGTCTCGAACTCCCGACCTCAGGATGATC 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTTCTCCAAGTTGGTCAGCGCTGGTCTCGAACTCCCGACCTCAGGATGATC 240

Qy    241 CCTCCCGTCTCGGGCCTCCCAAAAGTGCTAGATTACAGGACTGGGCCACCACCCGCTGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCCGTCTCGGGCCTCCCAAAAGTGCTAGATTACAGGACTGGGCCACCACCCGCTGCC 300

Qy    301 TGGCTAATTTTTGTGGGTAGAAAACAGGGTTTCACTGATGCTCAGCTGGTCTCCTGAGC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCTAATTTTTGTGGGTAGAAAACAGGGTTTCACTGATGCTCAGCTGGTCTCCTGAGC 360

Qy    361 TCAAGCAGTCAACCTGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TCAAGCAGTCAACCTGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC 420

Qy    421 CTGGGCCTTTTTATTTTATTTTTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGGCCTTTTTATTTTATTTTTTTTTAAGACACAGGTGTCCCACTCTTACCCAGGATGAAG 480

Qy    481 TCCAGTGGTGTGATCACAGCTCACTGCAGCCTCGAGTCAACTCCTGAGATCAAGCAATCCTCCTG 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TCCAGTGGTGTGATCACAGCTCACTGCAGCCTCGAGTCAACTCCTGAGATCAAGCAATCCTCCTG 540

Qy    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACTACACCTGGCTAATTTTTA 600

Qy    601 TTTTTATTTTAATTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTTTTATTTTAATTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 660

Qy    661 GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCGCCCCA 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCGCCCCA 720

Qy    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTGTATTTTT 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTGTATTTTT 780

Qy    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGTGATC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGTGATC 840

Qy    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCTTA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCTTA 900

Qy    901 TTTTTAATTTTTGTTTGTTTGTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTTTAATTTTTGTTTGTTTGTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAT 960
```

```
Qy    961 GGCCAAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGTCTCA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GGCCAAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGTCTCA 1020

Qy   1021 GCCTCCCAAGCAGCTGGGATTACAGGCGCCACCACCACCCGGCTCAAGCGATTCTCCTGTATTT 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCCTCCCAAGCAGCTGGGATTACAGGCGCCACCACCACCCGGCTCAAGCGATTCTCCTGTATTT 1080

Qy   1081 TCATTAGAGACGGGGTTTCACCATGTTGCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TCATTAGAGACGGGGTTTCACCATGTTGCAGGCTGGTCTCAAACTCCTGACCTCAGGT 1140

Qy   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCCAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCCAG 1200

Qy   1201 CCGGCTAATTTTTAGATAAAAAAAATATGTAGCAATGGGGGTCTTGCTATGTTGCCCAGGCT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CCGGCTAATTTTTAGATAAAAAAAATATGTAGCAATGGGGGTCTTGCTATGTTGCCCAGGCT 1260

Qy   1261 GGTCTCAAACTTCTGGCTTCAATCCTTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTCA 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GGTCTCAAACTTCTGGCTTCAATCCTTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTCA 1320

Qy   1321 CATTTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATTACAATAAACATGT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CATTTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATTACAATAAACATGT 1380

Qy   1381 CAAACCTGCAAATTCAGTAGTGTAACAGAGTTCTTTATTAAACTTTTAAACAAAGCTTTAGAG 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CAAACCTGCAAATTCAGTAGTGTAACAGAGTTCTTTATTAAACTTTTAAACAAAGCTTTAGAG 1440

Qy   1441 CA 1442
          ||
Db   1441 CA 1442
```

```
RESULT 2
LOCUS       AR072690                1442 bp    DNA        linear      PAT  28-AUG-2000
DEFINITION  Sequence 120 from patent US 5948634.
ACCESSION   AR072690
VERSION     AR072690.1  GI:9999454
KEYWORDS
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1442)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of alzheimer's
            disease
  JOURNAL   Patent: US 5948634-A 120 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="unknown"
ORIGIN
BASE COUNT    314 a    402 c    311 g    415 t

Query Match    100.0%;  Score 1442;  DB 6;  Length 1442;
Best Local Similarity  100.0%;  Pred. No. 0;  Mismatches 0;  Indels 0;  Gaps 0;
Matches 1442;  Conservative 0;

Qy      1 TTTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGGCAA 60

Qy     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy    121 CAGTAGCTGGGATTACAGGCGCATGCCACCACGCGCCGGCTAATTTTTGTATTTTTTAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

RESULT 3

```
LOCUS       AR073235                1442 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 120 from patent US 5948888.
ACCESSION   AR073235
VERSION     AR073235.1  GI:9999998
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1442)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5948888-A 120 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="unknown"
BASE COUNT      314 a    402 c    311 g    415 t
ORIGIN
```

Query Match   100.0%; Score 1442; DB 6; Length 1442;
Best Local Similarity   100.0%; Pred. No. 0;
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60

Qy     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy    121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG 180

Qy    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATC 240

Qy    241 CCTCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CCTCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCC 300

Qy    301 TGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGGTCTC 360

Qy    361 TAGAGCAGTCACCAACTCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGGTGC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TAGAGCAGTCACCAACTCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGGTGC 420

Qy    421 CTGGCCTTTTATTTTTATTTTTTTTTTTAAGACACAGGGTCTCACTCTGTCACCCAGGATGAAG 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGGCCTTTTATTTTTATTTTTTTTTTTAAGACACAGGGTCTCACTCTGTCACCCAGGATGAAG 480
```

RESULT 4
AF010144

```
LOCUS       AF010144                1442 bp    mRNA    linear   PRI 02-NOV-2001
DEFINITION  Homo sapiens neuronal thread protein AD7c-NTP mRNA, complete cds.
ACCESSION   AF010144
VERSION     AF010144.1  GI:3002526
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1442)
  AUTHORS   Monte,S.M., Ghanbari,K., Frey,W.H., Beheshti,I., Averback,P.,
            Hauser,S.L., Ghanbari,H.A. and Wands,J.R.
  TITLE     Characterization of the AD7C-NTP cDNA expression in Alzheimer's
            disease and measurement of a 41-kD protein in cerebrospinal fluid
  JOURNAL   J. Clin. Invest. 100 (12), 3093-3104 (1997)
  MEDLINE   98064067
   PUBMED   9399956
REFERENCE   2  (bases 1 to 1442)
  AUTHORS   de la Monte,S.M. and Wands,J.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUN-1997) MGH Cancer Center, Massachusetts General
            Hospital, 149 13th Street, Room 7308, Charlestown, MA 02129, USA
FEATURES             Location/Qualifiers
     source          1..1442
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /tissue_type="neuronal"
     CDS             15..1142
                     /note="contains expressed Alu-containing domains; 41 kDa
                     phosphoprotein; cell death and neuritic sprouting observed
                     in neuronal cells transfected with the AD7C-NTP cDNA"
                     /codon_start=1
                     /product="neuronal thread protein AD7c-NTP"
                     /protein_id="AAC08737.1"
                     /db_xref="GI:3002527"
                     /translation="MEFSLLLPRLECNGAISAHRNLRLPGSSDSPASASPVAGITGMC
                     THARLLYFFLVEMEFLHVGQAGLELPTSDDPSVSASQSARYRTGHHARLCLANFCGR
                     NRVSLMCPSWSPELKQSTCLSLPKCWDYRRAAVPGLFLFFLRHRCPLTQDEVQWCD
                     HSSLQPSTPEIKHPPASASQVAGTKDMHHYIWLIFIFNFLRQSLNSVTQAGVQWRN
                     LGSLQPLPPGFKRLFSCPSLLSSWDYRRPRLANFFVFLVEMGFTMPARLILISGPCDL
                     PASASQSAGITGVSHHARLIFNFCLFEMESHSVTQAGVQWPNLGSLQPLPGLKRFSC
                     LSLPSSWDYGHLPPHPANFCIFIRGVSPYLSGWSQTPDLR"
ORIGIN

BASE COUNT      314 a    402 c    311 g    415 t
```

```
Query Match          100.0%;   Score 1442;   DB 9;   Length 1442;
Best Local Similarity   100.0%;
Matches 1442;   Conservative  0;   Mismatches  0;   Indels  0;   Gaps  0;
                                  Pred. No. 0;
```

```
Qy     1 TTTTTTTTTTTGAGATGGGAGTTTTTCGCCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TTTTTTTTTTTGAGATGGGAGTTTTTCGCCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAA 60

Qy    61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 120

Qy   121 CAGTAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CAGTAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG 180

Qy   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 TAGAGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240

Qy   241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCACCATGCCCGGCC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CCTCCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACCACCATGCCCGGCC 300

Qy   301 TGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 TGGCTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTGGTCTCCTGAGC 360

Db   421 CTGGCCTTTTTTATTTTATTTTTTAAGACACAGGTGTCCACTCTTACCCAGGATGAAG 480

Qy   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAATCCTCTG 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 TGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAACTCCTGAGATCAAGCAATCCTCTG 540

Qy   541 CCTCAGCCTCCCAAGTAGCTGGGACCACAAAAGACATCACCACCACTGGGTTAATTTTTA 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 CCTCAGCCTCCCAAGTAGCTGGGACCACAAAAGACATCACCACCACTGGGTTAATTTTTA 600

Qy   601 TTTTTTATTTTTAATTTTTTGAGACAGAGTCTCACCCACCACCACAAGCTGGAGTGCAGT 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTTTTTATTTTTAATTTTTTGAGACAGAGTCTCACCCACCACCACAAGCTGGAGTGCAGT 660

Qy   661 GGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 GGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCA 720

Qy   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCACCACCACCCCACCCTAGCTAATTTTTTTTTTT 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GCCTCCTGAGTAGCTGGGACTACAGGCGCACCACCACCCCACCCTAGCTAATTTTTTTTTTT 780

Qy   781 TAGTAGAGATGGGGGGTTCACCATGTTCGCCAGGCTGGTCTCGAACTCCTGACCTTGTGATC 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TAGTAGAGATGGGGGGTTCACCATGTTCGCCAGGCTGGTCTCGAACTCCTGACCTTGTGATC 840

Qy   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCTTA 900
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCTTA 900

Qy   901 TTTTTTAATTTTTGTTTGTTTGTTTGAAAATGGAGTCTCACTCTGTTACCCAGGCTGGCAAT 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTTTTAATTTTTGTTTGTTTGTTTGAAAATGGAGTCTCACTCTGTTACCCAGGCTGGCAAT 960

Qy   961 GGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GGCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020

Qy  1021 GCCTCCCAAGCAGCTGGGATTACGGGCGCACCACCACACCCGGCTAATTTTTTGTATTT 1080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 GCCTCCCAAGCAGCTGGGATTACGGGCGCACCACCACACCCGGCTAATTTTTTGTATTT 1080

Qy  1081 TCATTAGAGACGGGGTTTCACCATATTTGTCAGGCTGGTCTCGAAACTCCTGACCTCAGGT 1140
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TCATTAGAGACGGGGTTTCACCATATTTGTCAGGCTGGTCTCGAAACTCCTGACCTCAGGT 1140

Qy  1141 GACCCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCTCACCCAG 1200
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 GACCCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCTCACCCAG 1200

Qy  1201 CCGGCTAATTTGAGTAAAAAAAAATATGTAGAGCAATGGGGTCTTGCTATGTTGCCCAGGCT 1260
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 CCGGCTAATTTGAGTAAAAAAAAATATGTAGAGCAATGGGGTCTTGCTATGTTGCCCAGGCT 1260

Qy  1261 GGTCTCAAACTTCTGGCTCATGCAATCCTTCCAAATGAGCCACAACAACCCAGCCAGTCA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GGTCTCAAACTTCTGGCTCATGCAATCCTTCCAAATGAGCCACAACAACCCAGCCAGTCA 1320

Qy  1321 CATTTTTTAAACAGTTGAACATCTTTAGTTATTACTAGAAAAGTAATAACAATATAAACATGT 1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CATTTTTTAAACAGTTGAACATCTTTAGTTATTACTAGAAAAGTAATAACAATATAAACATGT 1380

Qy  1381 CAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTATATACTAGAAAAAGCTTTTAGAG 1440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 CAAACCTGCAAATTCAGTAGTAACAGAGTTCTTTTATATACTAGAAAAAGCTTTTAGAG 1440

Qy  1441 CA 1442
         ||
Db  1441 CA 1442
```

```
Qy    361  TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGCAGCCGTGC  420

Qy    421  CTGGCCTTTTATTTTATTTTTTTATTTTTTTAAGACACAGGGTGTCCCACTCTTACCCTG  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  CTGGCCTTTTATTTTATTTTTTTATTTTTTTAAGACACAGGGTGTCCCACTCTTACCCTG  480

Qy    481  TGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCACTCCCTGGGCTCAAGCATCTCCTG   540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  TGCAGTGGTGTGATCACAGCTCACTGCAGCCTCCACTCCCTGAGATCAAGCATCCTCCTG  540

Qy    541  CCTCAGCCTCCCAAGTAGCTGGGACCACAAAGACATGCACCACCACTACACCTGGCTAAT  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  CCTCAGCCTCCCAAGTAGCTGGGACCACAAAGACATGCACCACCACTACACCTGGCTAAT  600

Qy    601  TTTTTATTTTTAATTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TTTTTATTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGT  660

Qy    661  GGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTATTCTCCTGCCCCA  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  GGGCGAATCGTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTGATTATTCTCCTGCCCA  720

Qy    721  GCCTCCTGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTGTATTTTT  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  GCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTTGTATTTTT  780

Qy    781  TAGTAGAGAATGGGGTTTCACCATGGTTCGCCAGGCTGGTCTCGAACTCTGTGATC        840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGTTGATCTCGAACTCTGGACCTTGTGATC  840

Qy    841  TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCGCCGGCTTA   900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCCGGCCTTA    900

Qy    961  GGCCAAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  GGCCAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA  1020

Qy   1021  GCCTCCCAAGCAGCAGCTGGGATTACGGGCCACCCACCACCACCGCCCGGCTAATTTTTT  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021  GCCTCCCAAGCAGCTGGGATTACGGGCCACCCACCACCACCGCCCGGCTAATTTTTGTATTT 1080

Qy   1081  TCATTAGAGACGGGGTTTCACCATATTTGTCAGGCTGGTCTCGAACTCCTGACCTCAGGT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081  TCATTAGAGACGGGGTTTCACCATATTGTAGCAATGGGGTCTCAAACTCCTGACCTCAGGT  1140

Qy   1141  GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCCACCCAG  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141  GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCCACCCAG  1200

Qy   1201  CCGGCTAATTAGATAAAAAATATAGTAGCAATGGGGTCTTGCTATGTTGCCCAGGCT    1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201  CCGGCTAATTTGAGTAAAAAATATGTTAGCAATGGGGTCTTGCTATGTTGCCCAGGCT    1260

Qy   1261  GGTCTCAAACTTCTGGCTTCATGCAATCCTCCAAATGAGCCACACCAGCCAGTCA        1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261  GGTCTCAAACTTCTGGCTTCATGCAATCCTCCAAATGAGCCACACACCAGCCAGTCA    1320

Qy   1321  CATTTTTTAAACAGTTACATCTTTATTTTAGTATATACTAGAAAGTAATAACATAAAACATGT  1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321  CATTTTTTAAACAGTTACATCTTTATTTTAGTATACTAGAAAGTAATAACATAAAACATGT  1380

Qy   1381  CAAAACCTGCAAAATTCAGTAGTAAACAGAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG  1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381  CAAAACCTGCAAATTCAGTAGTAACAGAGAGTTCTTTTATAACTTTTAAACAAAGCTTTAGAG  1440

Qy   1441  CA 1442
           ||
Db   1441  CA 1442
```

```
Db   1441  CA 1442

RESULT 5
HS886K2/c
LOCUS       HS886K2        124001 bp    DNA    linear    PRI 23-NOV-1999
DEFINITION  Human DNA sequence from clone 886K2 on chromosome 1p35.1-36.12
            Contains elongin A (RNA polymerase II elongation factor),
            hydroxymethylglutaryl-CoA lyase, GALE (UDP-galactose-4-epimerase)
            ESTs, GSS, STS, CpG island, complete sequence.
ACCESSION   AL031295
VERSION     AL031295.1  GI:4376011
KEYWORDS    HTG; CpG island; elongin A; HMGCL; hydroxymethylglutaryl-CoA
            lyase; RNA polymerase II elongation factor.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 124001)
  AUTHORS   Frankland,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-APR-1999) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 7, 1999 this sequence version replaced gi:4056542.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence is the entire insert of clone 886K2. This sequence
            has been finished according to sequence map criteria as follows. An
            attempt is made to resolve all sequencing problems, such as
            compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure,
            feature key.
            This sequence was generated from part of bacterial clone contigs of
            human chromosome 1, constructed by the Sanger Centre Chromosome 1
            Mapping Group. Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            886K2 is from the library RPCI5 constructed at the Roswell Park
            Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/ VECTOR: pCYPAC2.
FEATURES             Location/Qualifiers
     source          1..124001
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="p35.1-36.12"
                     /clone="RP5-886K2"
                     /clone_lib="RPCI-5"
     repeat_region   50..160
                     /note="AluY repeat: matches 195..302 of consensus"
     repeat_region   161..324
                     /note="L1MA5 repeat: matches 6133..6297 of consensus"
     repeat_region   327..594
                     /note="AluJb repeat: matches 1..294 of consensus"
     repeat_region   597..863
                     /note="AluJo repeat: matches 5..298 of consensus"
     prim_transcript 868..1170
                     /note="AluYb8 repeat: matches 1..310 of consensus"
     prim_transcript <1171..>1670
                     /note="match: multiple ESTs"
     repeat_region   1671..1848
                     /note="match: multiple ESTs"
     prim_transcript 1848
                     /note="MIR repeat: matches 67..261 of consensus"
     repeat_region   complement(<1849..>2287)
                     /note="match: multiple ESTs"
     repeat_region   2288..2429
                     /note="match: AA694386 AA234657 W84861 AA034271 W84860"
                     /note="match: A1276680 AA525120 AA032081 A1275142"
                     /note="match: AA232246 ...
```

repeat_region   2430..2714
                /note="AluJo/FRAM repeat: matches 157..300 of consensus"
repeat_region   2728..3063
                /note="AluSq repeat: matches 1..287 of consensus"
repeat_region   3385..3686
                /note="LIM4 repeat: matches 2200..2535 of consensus"
repeat_region   3709..4050
                /note="AluSx repeat: matches 1..310 of consensus"
repeat_region   4064..4212
                /note="THE1C repeat: matches 1..371 of consensus"
repeat_region   4358..4506
                /note="FRAM repeat: matches 1..155 of consensus"
repeat_region   4788..4878
                /note="MER5A repeat: matches 15..182 of consensus"
repeat_region   4879..5178
                /note="AluJo repeat: matches 54..149 of consensus"
repeat_region   5181..5484
                /note="AluSx repeat: matches 1..302 of consensus"
repeat_region   5485..5653
                /note="AluSg repeat: matches 1..305 of consensus"
repeat_region   5654..5788
                /note="AluJo repeat: matches 149..304 of consensus"
repeat_region   5789..6089
                /note="AluSx repeat: matches 1..132 of consensus"
repeat_region   6090..6260
                /note="AluSp repeat: matches 1..297 of consensus"
repeat_region   6282..6580
                /note="AluSx repeat: matches 132..312 of consensus"
repeat_region   6596..7224
                /note="AluY repeat: matches 1..296 of consensus"
repeat_region   7225..7293
                /note="L2 repeat: matches 1995..2748 of consensus"
repeat_region   7294..7753
                /note="AluJb repeat: matches 152..220 of consensus"
repeat_region   7756..8048
                /note="L2 repeat: matches 1543..1995 of consensus"
repeat_region   8050..8358
                /note="AluSg repeat: matches 1..279 of consensus"
repeat_region   8420..8987
                /note="AluY repeat: matches 1..309 of consensus"
repeat_region   8998..9090
                /note="L2 repeat: matches 873..1539 of consensus"
repeat_region   9060..9105
                /note="L2 repeat: matches 1543..1627 of consensus"
repeat_region   9180..9279
                /note="L2 repeat: matches 827..872 of consensus"
repeat_region   9489..9671
                /note="MIR repeat: matches 35..139 of consensus"
misc_feature   10046..10047
                /note="MER5A repeat: matches 1..189 of consensus"
repeat_region   10030..10594
                /note="Region of tandem repeat. Region of tandem repeat each element
                48 base pairs and typical sequence:
                GAGGAACTGTGTTAGATTATATACAATCCCTTCTGGATTC AGGTTCGT
                Restriction digest information (HINDIII) suggests 3 copies
                of the repeat are missing from this assembly."
repeat_region   10594..10678
                /note="L2 repeat: matches 2676..2750 of consensus"
repeat_region   11227..11408
                /note="MIR repeat: matches 12..87 of consensus"
repeat_region   11470..11776
                /note="MIR repeat: matches 20..232 of consensus"
repeat_region   12244..12546
                /note="AluY repeat: matches 1..303 of consensus"
repeat_region   13021..13318
                /note="AluSq repeat: matches 1..311 of consensus"
repeat_region   13319..13389
                /note="AluSx repeat: matches 1..295 of consensus"
repeat_region   13390..13685
                /note="MER5A repeat: matches 73..137 of consensus"
repeat_region   13686..13730
                /note="MER5A repeat: matches 27..73 of consensus"

repeat_region   13759..13823
                /note="L2 repeat: matches 2146..2215 of consensus"
repeat_region   13824..13959
                /note="L2 repeat: matches 2012..2146 of consensus"
repeat_region   13960..14115
                /note="FLAM_A repeat: matches 1..133 of consensus"
repeat_region   14221..14479
                /note="L2 repeat: matches 2146..2146 of consensus"
repeat_region   14660..14709
                /note="MER21B repeat: matches 156..456 of consensus"
repeat_region   14953..15263
                /note="MIR repeat: matches 49..98 of consensus"
repeat_region   15508..15884
                /note="AluSx repeat: matches 1..311 of consensus"
repeat_region   15724..15857
                /note="MIR repeat: matches 2..218 of consensus"
repeat_region   15961..16086
                /note="FLAM_A repeat: matches 1..133 of consensus"
repeat_region   16087..16394
                /note="FAM repeat: matches 12..163 of consensus"
repeat_region   16395..16409
                /note="AluSg repeat: matches 3..310 of consensus"
repeat_region   16722..16804
                /note="FAM repeat: matches 163..174 of consensus"
repeat_region   16908..17483
                /note="L2 repeat: matches 2582..2691 of consensus"
repeat_region   17679..17881
                /note="L1M3 repeat: matches 5598..6185 of consensus"
repeat_region   18516..18927
                /note="L2 repeat: matches 2285..2490 of consensus"
repeat_region   18928..19043
                /note="L2 repeat: matches 2273..2646 of consensus"
repeat_region   19044..19149
                /note="FLAM_C repeat: matches 1..116 of consensus"
repeat_region   19750..20088
                /note="MLT1D repeat: matches 114..505 of consensus"
repeat_region   20092..20153
                /note="31 copies 2 mer aa 81% conserved"
repeat_region   20170..20479
                /note="MLT1D repeat: matches 6..114 of consensus"
repeat_region   20499..20633
                /note="L2 repeat: matches 2623..2750 of consensus"
repeat_region   20639..20946
                /note="AluSx repeat: matches 3..312 of consensus"
repeat_region   20959..21232
                /note="L2 repeat: matches 1811..2198 of consensus"
repeat_region   21240..21532
                /note="AluJb repeat: matches 1..308 of consensus"
repeat_region   21534..21837
                /note="AluSx repeat: matches 1..308 of consensus"
repeat_region   21839..21936
                /note="AluSg repeat: matches 1..293 of consensus"
repeat_region   21952..22283
                /note="MER5B repeat: matches 84..175 of consensus"
repeat_region   22303..22605
                /note="AluSp repeat: matches 2..310 of consensus"
repeat_region   22665..22969
                /note="AluSg repeat: matches 1..304 of consensus"
repeat_region   23308..23605
                /note="AluSp repeat: matches 1..298 of consensus"

Query Match
Best Local Similarity 90.2%; Conservative 97.4%; Pred. No. 0; Mismatches 26; Indels 12; Gaps 8;
Matches 1409; Score 1301.4; DB 9; Length 124001;

```
Qy      1 TTTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCCAA   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6086 TTTTTTTTTTTTTTTTGAGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCCAA 6027

Qy     61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   6026 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 5967
```

```
RESULT 6
LOCUS       AR051479                1381 bp    DNA        linear   PAT 29-SEP-1999
DEFINITION  Sequence 49 from patent US 5830670.
ACCESSION   AR051479
VERSION     AR051479.1  GI:5974843
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1381)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5830670-A 49 03-NOV-1998;
FEATURES             Location/Qualifiers
     source          1..1381
                     /organism="unknown"
BASE COUNT      296 a    384 c    302 g    399 t
ORIGIN

Query Match   74.9%;  Score 1080.2;  DB 6;  Length 1381;
Best Local Similarity  94.4%;  Pred. No. 0;
Matches 1317;  Conservative 0;  Mismatches 53;  Indels 25;  Gaps 18;
```

```
Qy   121 CAGTAGCTGGGATTACAGGCATGNGCCACCACGGCTCGGCTAATTTTTGTATTTTTTAG  180
         |||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  5966 TAGTAGCTGGGATTACAGGCATGTGCCACCACGGCTCGGCTAATTTTTGTATTTTTTAG 5907

Qy   181 TAGAGATGGGAGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCGGACCTCAGATGATC  240
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5906 TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGGACCTCAGATGATC 5847

Qy   241 CCTCCGTCTCGGCCTCCCAAAGTGCT---AGATACAGGACTGGCCACCATGCCGG-CTC  296
         ||||||||||||||||||||||||||      |||||||||||||||||||||||  |||
Db  5846 CTCCCGTCTCGGCCTCCCAAAGTGCTGAGATTACAGGACTGGCCACCATGCCGCCTC 5787

Qy   297 TGCCTGGCTAATTTTTGTGGTTAGTAGAAACAGGGTTTCACTGATG-TGCCCAGGCTGGTCTCC  355
         ||||||||||||||||||||| |||||||||||||||||||||||  ||||||||||||||||
Db  5786 TGCCTGGCTAATTTTTGTGTTAGTAGAGACAGGGTTTCACTGATGTTGCCCAGGCTGGTCTCC 5727

Qy   356 TGAGCTCAAGAGAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC  413
         |||||||||||||||||||||||||||||||||||||||||||||||||||  ||| ||
Db  5726 TGAGCTCAAGAGAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC 5667

Qy   414 GCCGGTGCCTGGCCTTTTTTATTTTTTATTTTTTTAAGACACAGGGTGTCCCACTCTTACCCAG  473
         |||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db  5666 ACCGGTGCCTGGCCTTTTTTATTTTTTATTTTTTTAAGACAGGGTGTCCCACTCTTACCCAG 5607

Qy   474 GATGAGAGTGCAGTGGGTGTGATCACAGCTCACTGCCAGCCTTCAAACTCCTGAGATCAAG-A  532
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  |
Db  5606 GATGAAGTGCAGTGGGTGTGATCACAGCTCACTGCCAGCCTTCAAACTCCTGAGATCAAGCAA 5547

Qy   533 TCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACCACTACACCTGGCT  592
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5606 TCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCCACCACTACACCTGGCT 5547

Qy   593 AATTTTTTATTTTTTATTTTTTTTGAGACAGAGTCTCAACTCGTCACCCAGGCTGG  652
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5546 AATTTTTTATTTTTTATTTTTTTTGAGACAGAGTCTC-ACTCTGTCACCCAGGCTGG 5487

Qy   653 AGTGCAGTGGCGCCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTC  712
         ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  5486 AATTTTTTATTTTTTATTTTTTGAGACGAGTCTC-ACTCTGTCACCCAGGCTGG 5428

Qy   713 CTGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTT  772
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5427 AGTGCAGTGGCGCCAATCTTGGCTCACTGCAACCTCCGGGTTCAAGTTATTCTC 5368

Qy   773 TGTATTTTTAGTAGAGATGGGG-TTCACCATGGTTGGCCAGGTTGGTCTCGATCTCTGGAC  831
         ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db  5367 CTGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACGCCTAGCTAATTTTTT 5308

Qy   832 CTTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG  891
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5307 TGTATTTTTAGTAGAGATGGGGTTTCACCATGGTTGGCCAGGTTGGTCTCGATCTCTGAC 5248

Qy   892 CCCGGCCTTATTTTTTTAATTTTTGTTTGTTTGTTTGAGAATCACTCACTCTGTTACCCAGG  951
         |||||||||||| |||||||||||||||||| ||||||||||||||||||||||||||
Db  5247 CTTGTGATCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG 5188

Qy   952 GAGTGCAAATGGCCAAAATCTCGGCGTCACTGCAACCTCTGCCTCCCGGGCTCAAGCGATTCT 1011
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5187 CCCGGCCTTATTTTTTTAATTTTTTTGTTTTGTTTGAATCTCACTCTGTTACCCAGG 5128

Qy  1012 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACCACACCCCCGGCCTAATTT 1071
         |||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  5127 GAGTGCAATGGCACAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATGAT 5068

Qy  1072 TTTGTTATTTTTTCATTTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTCTCAAACTCCTG 1131
         ||||||||||||||||| |||||||||| |||||||||||||||||||||||||||||||||
Db  5067 CCTGTCTCAGCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACCACGTCCCGGCTA--ACACCCCCCTAATT 5010

Qy  1132 ACCTCAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 1191
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5009 TTTGTATTTTTCATTTAGAGCGGGGTTTCACCATATTTGGTAGAAACAGGGTTTCACTGCAAC 4950

Qy  1192 CTCACCCAGGCGGCTAATTTAGATAAAAAATATGTAGCAAATGGGGGGGTCTTGCTATGTT 1251
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4949 ACCTCAGGTGACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCAC 4890

Qy  1252 GCCCAGGCTGGTCTCAAACTCTGGCTCAAATGAGGACCACCAGTGCCGGG--CT 1311
         |||||||||||||||||||||||| ||||||||||||||||||||||||  ||
Db  4889 CTCACCCAGGCGGCTAATTTAGATAAAAAATATGTAGCAAATGAGGGGGTCTTGCTATGTT 4830

Qy  1312 AGCCAGTCACATTTTTTAAACAGTTACATCTTATTTTAGTGTAATAGAAGTAATACAA 1371
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4829 GCCCAGGCTGGTCTCAAACTCTGCCTCAAGCAGTGCTGGGATTACAGGCGTCCGCCACCATGCCCGGCC 4770

Qy  1372 TAAACATGTCAAACCTGCAAATTCAGTAGTAGTAGTGTCTTCTTTTATAATTTTAAACAAA 1431
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4769 AGCCAGTCACATTTTTTAAACAGTTACAICTTATTTTAGTGTAATAGAAGTAATACAA 4710

Qy  1432 GCTTTTAG 1438
         ||  |||
Db  4709 TAAACATGTCAAAACCTGCAAATTCAGTAGTAGTAGTGTCTTTTATAACTTTTAAACAAA 4650

Db  4649 GTATCAG 4643
```

```
RESULT 7
LOCUS       AR072619                1381 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 49 from patent US 5948634.
ACCESSION   AR072619
VERSION     AR072619.1  GI:9999383
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1381)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of alzheimer's
            disease
  JOURNAL   Patent: US 5948634-A 49 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1..1381
                     /organism="unknown"
BASE COUNT      296 a    384 c    302 g    399 t
ORIGIN

Query Match     74.9%;  Score 1080.2;  DB 6;  Length 1381;
Best Local Similarity   94.4%;  Pred. No. 0;
Matches 1317;  Conservative 0;  Mismatches 53;  Indels 25;  Gaps 18;
```

Qy   415 CCGTGCCTGCCTCAGCCTCCCTTTTTTATTTTTATTTTTAAGACACAGGTGTCCCAGG 474
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2 TTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTCCCAGGCTGGAGTGCAATGGCGCAAT 61

Qy   475 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCACCTCCTGAGATCTTACCCAGG 533
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC 121

Qy   534 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACACAGGTGTCCCAGGCTA 593
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   122 AGTAGCTGGGATTACAGGCGTGCACCACGCTCCCGGCTAATTTTGTATTTTTAGT 181

Qy   594 ATTTTTATTTTTATTTTTTAATTTTTTGAGACAGAGTCTC 653
             ||||||||||||||||||||||||||||||||
Db   182 AGAGATGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTC 239

Qy   654 GTGCAGTGGCGCAATCTCTTGGCTCACTGCAACCTCCG 713
             ||||||||||||||||||||||||||||||||||||
Db   240 CTGAGCTCAAGCAATCCACCTGCCTCAGCCTCCCAAAGTGCT 299

Qy   714 TGCCCCAGGCCTCCTGAGTAGCTGGGACTACAGG 773
             |||||||||||||||||||||||||||||||||
Db   296 CTGCCTGGCTAATTTTTGTGGTAGAAACAGGGTT 359

Qy   774 GTATTTTTTAGTAGAGATGGGG 832
             ||||||||||||||||||||
Db   360 CTGAGCTCAAGCAGTGCCACCTGCCTCGGCCTCCCAAAGTGCTGG 418

Qy   833 TTGTGATCTGTGCCTGCCTCGGCCT 891
             |||||||||||||||||||||||
Db   419 CCGTGCCTGGCCTTTTTTATTTTTT 474

Qy   892 CCCGGCCTTATTTTTAATTTTTGT 951
             ||||||||||||||||||||||
Db   415 CCGTGCCTGCCTCAGCCTCCCTTTTTTATTTTTATTTTTAAGACACAGGTGTCCCAGG 474

Qy   952 GAGTGCAATGGCCAAAATCTCGGCTCACTGC 1011
             |||||||||||||||||||||||||||||
Db   475 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCACCTCCTGAGATCTTACCCAGG 533

Qy   1012 CCTGTCTCAGCCTCCCAAGAGCAGCTGGGAATTACGG 1071
             |||||||||||||||||||||||||||||||||||
Db   534 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGA 593

Qy   1072 TTTGTATTTTCATTAGAGGGGGG 1128
             ||||||||||||||||||||
Db   594 ATTTTTATTTTTATTTTTTAATTTTTTGAGACAGAGTCTC 653

Qy   1129 GACCTCAGGTGACCCACCTGCCTCCA 1188
             ||||||||||||||||||||||||
Db   654 GTGCAGTGGCGCAATCTCTTGGCTCACTGCAACCTCCG 713

Qy   1191 CCTCACCCAGCCGGCCTAATTAGATAAAAAAATATGT 1250
             ||||||||||||||||||||||||||||||||||
Db   714 TGCCCCAGGCCTCCTGAGTAGCTGGGACTACAGG 773

Qy   1189 CCTCACCCAGCCGGCCTAATTTGGAATAAAAAATGTAGCAACAATGGGGG 1246
             ||||||||||||||||||||||||||||||||||||||||||||||
Db   657 GTGCAGTGGCGCAATCTCTTGGCTCACTGCAACCTCCG 716

Qy   1247 TGCCCAGGCTGGTCTCTCAAAACTTCTGGTTCATGCAATCC 1306
             |||||||||||||||||||||||||||||||||||||||||
Db   539 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGA 597

Qy   1251 TGCCCAGGCTGGTCTCTCAAAACTTCTGGAATAAAAATGTAGCAACAACACC 1310
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   479 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCACCTCCTGAGATCAAGCAAT 538

Qy   1311 CAGCCAGTCACATTTTTAAACAGTTACATTCTTTTATTTTT 1370
             |||||||||||||||||||||||||||||||||||||||||
Db   534 CCTCCTGCCTCAGCCTCCCAAGTAGCTGGGA 593

Qy   1307 CAGCCAGTCACATTTTTTAAACAGTTACATTCTTTATTTTTAGTAATACA 1366
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   475 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCACCTCCTGAGATCTTACCCAGG 533

Qy   1371 ATAAACATGTCAAAC 1385
             |||||||||||||||
Db   419 CCGTGCCTGGCCTTTTTTATTTTTT 474

Qy   1367 ATAAACATGTCAAAC 1381
             |||||||||||||||
Db   415 CCGTGCCTGCCTCAGCCTCCCTTTTTTATTTTTATTTTTAAGACACAGGTGTCCCAGG 474

```
RESULT 8
AR073164

LOCUS       AR073164                1381 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 49 from patent US 5948888.
ACCESSION   AR073164
VERSION     AR073164.1  GI:999927
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1381)
  AUTHORS   de la Monte,S. and Wands,J.R.
  TITLE     Neural thread protein gene expression and detection of Alzheimer's
            disease
  JOURNAL   Patent: US 5948888-A 49 07-SEP-1999;
FEATURES             Location/Qualifiers
  source          1..1381
                  /organism="unknown"
BASE COUNT      296 a    384 c    302 g    399 t
ORIGIN
```

Query Match   74.9%; Score 1080.2; DB 6; Length 1381;
Best Local Similarity   94.4%;
Matches 1317; Conservative   0; Mismatches   53; Indels   25; Gaps   18;

```
Qy      2 TTTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCCAAT   61
          ||||||||| || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTTTTGAGACGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCGAT   60

Qy     62 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC   121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC   120

Qy    122 AGTAGCTGGGATTACAGGCGCATGCCACCACGCCCGGCTAATTTTTGTATTTTTAGT   181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGTAGCTGGGATTACAGGCATGCCA-CCACGCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCC   179

Qy    182 AGAGATGGAGTTT--CTCCATGTTGGTCAGGCTGGTCTCGAAACTCCCGACCTCAGATGAT   239
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||
Db    180 AGAGAGTGGAGTTTAACTCCATGTTGGTCAGGCTGGTCTCGAAACTCCCGACCTCAGATGAT   239

Qy    240 CCCTCCGTCTCGCCTCCCAAAGTGCTGGGATTACAGGCGT-CT   295
          ||||||||||||||||||||||||||||||||||||||||| ||
Db    240 CTCCCGTCTCGGCCGCTGCCCAAAGTGCTGGGATTACAGGCGTGCCAGGCCT   299

Qy    296 CTGGCCTGGCTAATTTTTTGTGGTAGAGACAGGGTTTCACTGGTG-TGCCCAAGCTGGTCTC   354
          ||||||||||||||||| ||||||||||||||||||||||||||| |||||||||||||||
Db    300 CTGCCTGGCTAATTTTTTGTATTTTTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGTCTC   359

Qy    355 CTGAGCTCAAGCAGTCCACCACGTCCCTCCCAAAGTGCTGGGATTACAGGCGTGAG   414
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 CTGAGCTCAAGCAGTGCAACCTCTGCCTCCCAAAGTGCTGGGATTACAGGCGT-CAG   418

Qy    415 CCGTGCCTGGCCTTTTTTATTTTTTATTTTTTTAAGACAGAGTGTCCCACTCTTACCCAG   474
          |||||||||||| ||||||||||||||||||| |||||||||||||||||||||||||
Db    419 CCGTGCCTGGCCTTTTTTTTTTTTTTTTGTAGAAACAGGGTTTCACTATGTTGCCCAGG   478

Qy    475 ATGAAGTGCAGTGGTGTGATCACGGCTCACTGCAGCCTCCGACCTCCTGAGATCAAGC-AT   533
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    479 ATGAAGTGCAGTGGTGTGATCACAGCTCACTGCAGCCTTCAAACCTCCTGAGATCAAGCGAAT   538

Qy    534 CCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACCAAAGACAATGCACCACCACCTGGCTA   593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    539 CCTCCCGCCTCAGCCTCCCAAGTAGCTGGGACCACAGGCGCCCACCACCACGCCCGGCTA   597

Qy    594 ATTTTTATTTTTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTCG-TA   653
          ||||||| ||||||||||||||||| ||||||||||||||||||||||||||||| ||
Db    598 ATTTTTTGTTTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGA-TCTC   656

Qy    654 GTGCAGTGGCGCGAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCC   713
          ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    657 GTGCAGTGGCGCAATCTTGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCC   716

Qy    714 TGCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAATTTTTT   773
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db    717 TGCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCAGCTAATTTTTT   776

Qy    774 GTATTTTTTAGTAGAGATGGGGG--TTCACCATGGTTGGCCAGGCTGGTCTCGATCTCTGGACC   832
          ||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||||
Db    777 GTATTTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCTTGACC   836

Qy    833 TTGTGATCTGCCTGCCTCGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCACG   891
          |||||||||||||||||||||||||| | ||||||||||||||||||||||||||||||
Db    837 TTGTGACTCGGCCCTGCCTCGGCCTACCCAAAGTGCTGGGATTACAGG--GTCGTGACTCACACG   894

Qy    892 CCCGGCCTATTTTTTAATTTTTGTTTGTTTGAGAATGGAATCTCACTCTGTTACCCAGGCTG   951
          ||||||| |||||||||||||| |||||||||||||||||||||||||||||||||||||
Db    895 GCCGGCCTATTTTTTAATTTTTTGTTTGTTTGAGAATCTCACTCTGTTACCCAGGCTG   954

Qy    952 GAGTGCAATGGCCAAATCTCGGCTCACTGCAACCTCTGCCTCAGCCTCCGGGGGCTCAAGCGATTCT   1011
          |||||||||||||| ||||||||||||||||||||||||||||||  ||||||||||||||||
Db    955 GAGTGCAATGG-CAAATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCT   1012

Qy   1012 CCTGTCTCAGCCTCCCAAGCACGCTGGGATTAGGGCACCTGCCACCACACACCCCCGGCTAATT   1071
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1013 CCTGTCTCAGCCTCCCAAGCAGCTGGGATTACAGG--ACCTGCCACCACACCCCCGGCTAATT   1070

Qy   1072 TTTGTATTTTCAATTAGAGGCGGGGTTTCAGCGCC--TTTTACCATATTGTCAGGCTGGGTCTCAAACTCCT   1130
          ||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||||||||
Db   1071 TTTGTATTTTCATTAGAGGCGGG--TTTACCATATTTGTCAGCTGGTCTCAAAACTCCT   1128

Qy   1131 GACCTCAGGTGACCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA   1190
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1129 GACCTCAGGTGACCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCA   1188

Qy   1191 CCTCACCCAGCCGCGGGCTAATTAGATAAAAAAAATATGTAGCAATGGG   1250
          ||||||||||| ||||||||| | ||||||||||||||| || |||
Db   1189 CCTCACCCAGCCGCTAATTGGAATAAAAATATGTAGCAATGTTGCAAATGGGGG   1246

Qy   1251 TGCCCAGGCTGGTCTCAAACTCTTGGGCTTCATGCGAATCCTTCCAAATGAGCCACAACAC   1310
          |||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db   1247 TGCCCAGGCTGGTCTCAAACTTCTGGGCTTCAAGTCAATCCTTCCAAATGAGCCACCACAC   1306

Qy   1311 CAGGCCAGTGCACATTTTTTAAACAGTTCACATCTTTATTTTATTACTAGAAAAGTAATACA   1370
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   1307 CAGGCCAGTGCACATTTTTTAAACAGTTCACATCTTTATTTTATTACTAGAAAGTAATACA   1366

Qy   1371 ATAAACATGTCAAAC   1385
          |||||||||||||||
Db   1367 ATAAACATGTCAAAC   1381
```

```
Qy    1131 GACCTCAGTGACCACTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1190
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1129 GACCTCAGTGACCACTGCCTCAGCCTTCCAAAGTGCTGGGATTACAGGCGTGAGCCA 1188

Qy    1191 CCTCACCCAGCCGGCTAATTTAGATAAAAAAATATGTAGCAAATGGGGTTTGCTATGT 1250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1189 CCTCACCCAGCCGGCTAATTTGGATAAAAAAATATGTAGCAAATGGGGTTTGCTATGT 1246

Qy    1251 TGCCCAGGCTGGTCTCAAACTTCTGGGCTTCATGCAAATCCTTCCCAAATGAGCCACAACACC 1310
           |||||||||||||||||||||||||||||||||||   ||||||||||||||||||||
Db    1247 TGCCCAGGCTGGTCTCAAACTTCTGGGCTTCAATCCTTCCCAAATGAGCCACAACACC 1306

Qy    1311 CAGGCCAGTCACATTTTTAAACAGTTACATCTTTTATTTTAGTATACTAGAAAGTAATACA 1370
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1307 CAGGCCAGTCACATTTTTTAAACAGTTACATCTTTTATTTTAGTATACTAGAAAGTAATACA 1366

Qy    1371 ATAAAACATGTCAAAC 1385
           ||||||||||||||||
Db    1367 ATAAACATGTCAAAC 1381
```

RESULT 9
AC087434
AC087433
LOCUS   AC087434   160714 bp   DNA   linear   HTG   05-JAN-2001
DEFINITION Pan troglodytes clone RP43-90F9, WORKING DRAFT SEQUENCE, 12
            unordered pieces.
ACCESSION   AC087434
VERSION   AC087434.1 GI:12039251
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Pan troglodytes.
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 160714)
AUTHORS   Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S.,
            Guan,X., Gupta,J., Ho,S.-L., Idol,J.R., Karlins,E., Lee-Lin,S.-Q.,
            Legaspi,R., Lim,M., Maduro,Q.L., Maduro,V.B., Masiello,C.,
            Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Prasad,A.,
            Shevchenko,Y., Snyder,B., Stantripop,S., Thomas,J.W., Thomas,P.J.,
            Tiongson,E.E., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A.,
            Wetherby,K.D., Zhang,L.-H. and Green,E.D.
TITLE   NISC Comparative Sequencing Initiative
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 160714)
AUTHORS   Green,E.D.
TITLE   Direct Submission
JOURNAL   Submitted (05-JAN-2001) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT   -------------- Genome Center
              Center: NIH Intramural Sequencing Center
              Center code: NISC
              Web site: http://www.nisc.nih.gov
              Contact: nisc_mouse@nhgri.nih.gov
            -------------- Project Information
              Center project name: aod
              Center clone name: 090F09
            -------------- Summary Statistics
              Sequencing vector: plasmid; n/a; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.990319
              Consensus quality: 152563 bases at least Q40
              Consensus quality: 154387 bases at least Q30
              Consensus quality: 155752 bases at least Q20
              Insert size: 137000; agarose-fp
              Insert size: 137000; pulse-field-gel
              Insert size: 159614; sum-of-contigs
              Quality coverage: 9.32x in Q20 bases; agarose-fp
              Quality coverage: 9.32x in Q20 bases; pulse-field-gel
              Quality coverage: 8.00x in Q20 bases; sum-of-contigs
            ----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *   1   2808: contig of 2808 bp in length
            *   2809   2908: gap of unknown length
            *   2909   6190: contig of 3282 bp in length
            *   6191   6290: gap of unknown length
            *   6291   12538: contig of 6248 bp in length
            * 12539   12638: gap of unknown length
            * 12639   20815: contig of 8177 bp in length
            * 20816   20915: gap of unknown length
            * 20916   28981: contig of 8066 bp in length
            * 28982   29081: gap of unknown length
            * 29082   40699: contig of 11618 bp in length
            * 40700   40799: gap of unknown length
            * 40800   49520: contig of 8721 bp in length
            * 49521   49620: gap of unknown length
            * 49621   62547: contig of 12927 bp in length
            * 62548   62647: gap of unknown length
            * 62648   76465: contig of 13818 bp in length
            * 76466   76565: gap of unknown length
            * 76566   98277: contig of 21712 bp in length
            * 98278   98377: gap of unknown length
            * 98378   127174: contig of 28797 bp in length
            * 127175   127274: gap of unknown length
            * 127275   160714: contig of 33440 bp in length.
FEATURES   Location/Qualifiers
     source   1..160714
                     /organism="Pan troglodytes"
                     /db_xref="taxon:9598"
                     /clone="RP43-90F9"
                     /clone_lib="RP43"
     misc_feature   1..2808
                     /note="assembly_fragment"
     misc_feature   2909..6190
                     /note="assembly_fragment"
     misc_feature   6291..12538
                     /note="assembly_fragment"
     misc_feature   12639..20815
                     /note="assembly_fragment"
     misc_feature   20916..28981
                     /note="assembly_fragment"
     misc_feature   29082..40699
                     /note="assembly_fragment"
     misc_feature   40800..49520
                     /note="assembly_fragment"
     misc_feature   49621..62547
                     /note="assembly_fragment"
     misc_feature   62648..76465
                     /note="assembly_fragment"
     misc_feature   76566..98277
                     /note="assembly_fragment"
     misc_feature   98378..127174
                     /note="assembly_fragment"
     misc_feature   127275..160714
                     /note="assembly_fragment"
                     /clone_end:SP6
                     /vector_side:right"
BASE COUNT   37810 a   40093 c   40289 g   40706 t   1116 others
ORIGIN

Qy   4 TTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTTTGCCCAGGCTGGAGTGCAATGCGCAATCT 63
           |||| ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 132561 TTTTGTTTGAGATGGAGTTTTT-GCTCTTGTTGTCCCAGGCTGAAGTGTAATGGCACGAGTCT 132619

Query Match   41.1%; Score 592.2; DB 2; Length 160714;
Best Local Similarity   71.2%; Pred. No. 6.5e-167;
Matches 944; Conservative 0; Mismatches 348; Indels 33; Gaps 11;

Qy   64 CAGCTCACCGGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG 123
```

```
Db 133692 AACTCCTAACTTCAGGTGATCACCGGCGTCCTTCGCCTTCCCAAAGTCGTGGGATTACAGGCA 133751
          ||||||| |||||| ||||| |||  |||| |||| ||||| ||||||||||||||||||||||
Qy   1184 TGAGGCACCTCACCCAGCC----GGCTAATTTAGATAATAAAAATATGTAGCAATGGGGG 1239

Db 133752 TGAGGCACCACACCCTGGCCTCTGCCCAACTAATTTTTAAAAAAATTTTGTAGACAGGA 133811
          ||||||||| ||||| |||| ||| |||  ||||||||| |||||||| || ||||||||
Qy   1240 TCTTGCCTATGTTGCCCAGGCTGGTCTCGAAACTTCTGGCCTTCATGCAATCCTTCCAAATGA 1299

Db 133812 TCTCACTATGTTGTCCAGACTGGGTCTTGAACTCCTGGGCTCAAGGGATACGTCGTCCTGTCTCA 133871
          |||| | |||||| ||||| |||| |||||| || || |||||||  ||| | ||| |||||||
Qy   1300 GCCAC 1304
          |||||
Db 133872 GCCTC 133876

Qy    244 CCGTCTCGGGCCTCCCAAAGTGCTGGGATTACAGGCGTTGGTCTTTTACAAGAAAACATTTA 303
          |||||| | |||||||||||||||| ||||||||||  ||| || |||||| | | ||||
Db 132796 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTTTACAAGAGTTTTGTTG 132854

Qy    184 AGATGGAGTTTCTCCATGTTGGTCAGCGCTGGTCTCGAACTCCCGACCTCAGGATGATCCCT 243
          |||||||||||||||||||||||| ||||||||||||||||| |||||| ||||||||||||
Db 132736 AGACGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGATGATCCGC 132795

Qy    124 TAGCTGGGATTACAGGCGTGCACCCACCGCGTCGGCTAATTTTTGTATTTTTTTTAGTAG 183
          |||| | ||||||||||  | | |||| |||||||||||| || | ||||||---AGT
Db 132680 TAGCTGGGATTACAGGCGTGTCACCACCGCGGCTAATTTTTGTATTTTT---AGT 132735

Qy    304 CTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGTCCCAAGCTGGTCTCTGAGCTCA 363
          ||||||||| | || ||| ||||||||| ||| ||| |||||||||||| ||||||||
Db 132855 CCCAGGATGAAGAGTGCAGTGCCATCTCCGGAATATGCCGGCCTCCAAACCAGCTGTTGGAGACTC 132914

Qy    364 AGCAGTGCACCTGTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTG----- 419
          |||| |||| ||| ||||||||| |||||||||||||||||||||||| |
Db 132915 CCCAGGCTGGAGTGCAGTGGTCACGATCTCGGCTCACTGCAAGCTCTGCCTCCCGGGTTCA 132974

Qy    420 -------CCTGGGCTTTTATTTATTTTAAGACACAGGTGTCCCACTCTTA 468
          |||||||||| |||| || | || |||||| || || ||||||||||||
Db 132975 CCCCCCATCCCATCTGCCTTTTTTTTTTTTTTTTGAGACAGA-GTCTTGTTTGTTG 133033

Qy    469 CCCAGGATGAAGAGTGCAGTGGTCACGATCTCGGCTCACTGCAAGCTCTGCTGAGATC- 527
          ||| || || |||| ||| |||||||||||||||||||||||| || || |  |
Db 133034 GCTGGGGTGCAATGGTGCAGTGGTCGCGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCA 133093

Qy    528 AAGCACATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCAACCACTACACC 587
          |||| || |||||| |||||||||||||| | |||||| | | | | ||| |  ||||||||
Db 133094 AAGAATCCTCCCACCTCAGCCTCCCAAATAGCTGGGACTATAGGTGCATGCCACCATGCC 133153

Qy    588 TGGCCTAATTTTTATTTTTATTTTTTGAGACAGTCAACTCTGTCACCCAG 647
          ||| ||||||||| ||||| |||||||||  ||| | |||| |||||||
Db 133154 CAACTAATTTTTTTTTTTTT--TTTTGAGATGGGAGTTTCGCTCTTGTTGCCCAG 133211

Qy    648 GCTGGAGTGCACTGGCGCCAATCTCGGCTCACTGCAACTTCTGCCTCCCAAGTTA 707
          ||||||||||| |||||| || |||||||||||||||| || ||||||||| ||
Db 133212 GCTGGGGTGCAATGGTGCAGTGGTCTCAACCTCTGGGTTCACACCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACA 133271

Qy    708 TTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCGCCACCACCGCCTAGTAAT 767
          ||||||| |||||||||||| ||||||||||| ||| | || | ||||| || | |
Db 133272 TTCTCTGCCTCAGCCTCCCCAAATAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCT 133331

Qy    768 TTTTTGTGATTTTTAGTAGAGATGGGG--TTCACCATGGTTGGCCAGGTTGGTCTCGATCTCTC 826
          |||||  | ||| ||||||||| || || |||||| ||||||||||| | ||||||| || |
Db 133332 AATTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAACTC 133391

Qy    827 TGGACCT---TGTGATCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGC 884
          ||| |||   |||| || | |||||||||||||||||||||||||||| ||| |||
Db 133392 CTGACCTCAGGAGATCCTCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGC 133451

Qy    885 CACCACGCCCGG-CTTATTTTTAATTTTTGTTTGTTTGAAATGAAATCACACCTGTTAC 943
          |||||| ||| |||| | |||||| |||| || || ||  | || ||||||||| ||
Db 133452 CACTGTGCCTGGCCTTTTTTTTTTTTTTTTGGAGACAAAGTCTCGTCTGTTGC 133511

Qy    944 CCAGGCTGGAGTGCAATGGCCAAATCTCGGCTCACTGCAACCTCCGGGCTCAA 1003
          ||||||||||||||| | | |||| |||||||||||||||| || | | ||||
Db 133512 CCAGGCTGGAGTGGAAATGGCACACACCTCGGCTCACTGCGAACCTCTGGCCTTTTGGGTTCAA 133571

Qy   1004 GCGATTCTCCTGTCTCAGCCTCCCGAGAGCTGGGATTACAGGCGCCACCACC 1063
          || |||||||| ||||||||||||| || ||||| | |||||| |||||||
Db 133572 GCGATTCTCCTGCCTCCCAAGTAGCTCCCAGGACTACAGGCCATGGCCACCACGCCC 133631

Qy   1064 CGCTAAATTTTGTATTTTTCATTAGAGGCGGGTTTCACCATATTTGTCAGGCTGGTCTCA 1123
          || ||||||||| ||||| || ||||| || |||||||||| ||||| |||||||||||
Db 133632 GGCTAATTTTTGTATTTTTAGTAGAGACGAGTTTCACCATGTTGGCCGGGCTGGTCTTG 133691

Qy   1124 AACTCCTGACCTTCAGGTGATCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG 1183
          ||||||| | ||||||||||||||| | ||| | |||||||||||||||||||||||||
```

from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see http://www.nhgri.nih.gov/DIR/GTB/CHR7 , send mailto:egreen@nhgri.nih.gov , or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
Clone CTB-52H6 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc.
VECTOR: pBeloBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is CTA-270D13. Actual start of this clone is at base position 135330 of CTA-270D13; actual end is at base position 99370 of CTB-52H6.

```
FEATURES             Location/Qualifiers
     source          1..99370
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q11.2-q22"
                     /clone="CTB-52H6"
                     /clone_lib="CTB-978SK-B"
     repeat_region   59..528
                     /rpt_family="MER2_type"
     repeat_region   529..837
                     /rpt_family="Alu"
     repeat_region   838..903
                     /rpt_family="MER2_type"
     repeat_region   1428..1729
                     /rpt_family="Alu"
     misc_feature    3190..3258
                     /note="match to EST AI208950 (NID:g3770892) qg60d12.x1"
     misc_feature    3193..3259
                     /note="similar to Mus musculus EST AA165787 (NID:g1744017)"
     misc_feature    3193..3259
                     /note="similar to Mus musculus EST W41925 (NID:g1325640)"
     misc_feature    3199..3224
                     /note="match to EST AA333593 (NID:g1985887) mc65g12.r1"
     misc_feature    3587..3684
                     /note="match to EST AA076195 (NID:g1616065) zm18a06.r1"
     misc_feature    3587..3684
                     /note="match to EST AL036193 (NID:g5927681)"
     misc_feature    3588..3684
                     /note="match to EST AA333593 (NID:g1985887)"
     misc_feature    3588..3684
                     /note="match to EST AI208950 (NID:g3770892) qg60d12.x1"
     misc_feature    3588..3684
                     /note="similar to Mus musculus EST AI208950 (NID:g3770892) qg60d12.x1"
     misc_feature    3588..3684
                     /note="similar to Mus musculus EST AA165787 (NID:g1744017) ms45b10.r1"
     misc_feature    3588..3684
                     /note="similar to Mus musculus EST W41925 (NID:g1325640) mc65g12.r1"
     misc_feature    3769..3833
                     /note="match to EST AL036193 (NID:g5927681)"
     misc_feature    3769..3833
                     /note="match to EST AA333593 (NID:g1985887) zm18a06.r1"
     misc_feature    3769..3833
                     /note="match to EST AA076195 (NID:g1616065) zm18a06.r1"
     misc_feature    3769..3833
                     /note="match to EST AL036193 (NID:g5927681)"

     misc_feature    3769..3833
                     /note="similar to Mus musculus EST W41925 (NID:g1325640) mc65g12.r1"
     misc_feature    3769..3826
                     /note="similar to Mus musculus EST AA165787 (NID:g1744017) ms45b10.r1"
     misc_feature    3769..3792
                     /note="match to EST AI208950 (NID:g3770892) qg60d12.x1"
     misc_feature    3770..3833
                     /note="match to EST AA852497 (NID:g2941090)"
     misc_feature    3785..3833
                     /note="match to EST R58433 (NID:g828491)"
     misc_feature    4726..4728
                     /note="match to EST AA076195 (NID:g1616065) zm18a06.r1"
     misc_feature    5708..5895
                     /note="match to EST AA852497 (NID:g2941090)"
     misc_feature    5716..5895
                     /note="match to EST AA333593 (NID:g1985887)"
     misc_feature    5708..5896
                     /note="match to EST D59085 (NID:g968719)"
     misc_feature    5731..5896
                     /note="similar to Mus musculus EST W28383 (NID:g1308331) mc65g12.r1"
     misc_feature    5708..5879
                     /note="match to EST R58433 (NID:g828491)"
     repeat_region   5708..5826
                     /rpt_family="Alu"
     repeat_region   6582..6695
                     /rpt_family="Alu"
     repeat_region   6696..6717
                     /rpt_family="(A)n"
     repeat_region   6924..7013
                     /rpt_family="MIR"
     repeat_region   7014..7228
                     /rpt_family="Alu"
     repeat_region   7279..7319
                     /rpt_family="MIR"
     repeat_region   7637..7677
                     /rpt_family="Alu"
     repeat_region   8214..8525
                     /rpt_family="T2_type"
     repeat_region   8526..8811
                     /rpt_family="Alu"
     repeat_region   8813..8866
                     /rpt_family="Alu"
     repeat_region   9022..9059
                     /rpt_family="A-rich"
     repeat_region   9729..9763
                     /rpt_family="(TATG)n"
     repeat_region   10803..10908
                     /rpt_family="(CACAT)n"
     repeat_region   11017..11137
                     /rpt_family="Alu"
     repeat_region   11492..11762
                     /rpt_family="Alu"
     repeat_region   11763..12059
                     /rpt_family="Alu"
     repeat_region   12060..12093
                     /rpt_family="Alu"
     repeat_region   12331..12470
                     /rpt_family="A-rich"
     repeat_region   12471..12597
                     /rpt_family="L1"
     repeat_region   12951..13298
                     /rpt_family="Alu"
     repeat_region   13299..13572
                     /rpt_family="MER1_type"
     repeat_region   13572..13572
                     /rpt_family="Alu"
```

Query Match 40.9%; Score 590.4; DB 9; Length 99370;

```
Best Local Similarity  70.8%; Conservative  0; Pred. No. 2.1e-166; Mismatches 346; Indels 43; Gaps 10;
Matches  941;

Qy     4 TTTTTTTTGAGATGGAGTGAGTTTTCGCTCTTGTTGTTGCCAGGCTGGAGTGCAATGGCGCAATCT 63
         |||| ||||||||||||  ||||| |||||||||||||||||||||||| ||||| |||||| | |||
Db 69541 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTTGCCAGGCTGAAGTGTAATGGCACAGATCT 69599

Qy    64 CAGCTCACCGCGAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCCAG 123
         ||||||||| | ||| |||||||||| |||||||||| ||||||| |||||||||||||| 
Db 69600 CAGCTCACCAACAACTTCCGCCTCCCAGGTTCAAGCAATTCTCCTACCTCAGCCTCCCGAG 69659

Qy   124 TAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTTTTAGTAG 183
         |||||||||||||||||||| |||||||||||||||||||||||||||||||   ||| 
Db 69660 TAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTT---AGT 69715

Qy   184 AGATGGAGTTTCTCCATGTTGGTGCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCCT 243
         ||||||||||||| ||||||||| |||||||||||||||||||||||| |||||| ||||| 
Db 69716 AGACGGGGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACCCCCGACCTCAGATGATCCGC 69775

Qy   244 CCGTCTCGGCCTCCCAAAGTGCTGTAGATACAGGACTACAGGCGCCCGGCTCTGCCTGG 303
         |||||||||||||||||||||||| || | | ||||||||| |||||||||||||||||
Db 69776 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTTACAGGAAAACATTTA 69834

Qy   304 CTAATTTTTGTGTGGTAGAAACAGGGTTTCACTGATGATGTGTCCTCCTGAGCTCA 363
         ||||||||| ||| ||||||||||||||||||||| |  | | ||||||||||||||
Db 69835 AGTATGAAAAGCACTACCTCCGGGAATATCGCACGGCCCCTCTGAAGACTC 69894

Qy   364 AGCAGTCCAACTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCGTGCCTG 423
         ||||||||| |||||| ||||| ||||||||||||||||||||| || |||  ||||||| 
Db 69895 TTCACTGTCTCTGGCTGGGCACCACCTGTCCACCTCAAAACTGGGCTGTGTCTCTT 69954

Qy   424 GCC-----------------TTTTTAATTTTATTTTTTTAAGACACAGGGTGTCCCACTC 465
         |||                 |||||||||||||||||||||||||||||||| |||||||
Db 69955 CCCCCCCCCCATCCCCTGCCTTTTGTGTGTGTGTGATCACAGCTCACTGCAGAGTCTTGTTTTTG 70014

Qy   466 TTACCCAGGATGAAGTGCAGTGCGTGGTTGATCACAGCTCACTGCAGCCTTCAAACTCCTGAGA 525
         |||||| ||||| |   ||||| |||||||||||||||||||||||||||| |||||||||| 
Db 70015 TTGCCCAGGCTGGAGTGCAGTGGCGCGAAACATGGCTCACTGCAGCCTTGATCTCCTGGGC 70074

Qy   526 TC-AAGCANTCCTCTGCCTCCAGCCTCCCAAGTAGCTGGGACCAAAGAACACAGCACTAC 584
         || |||| | |||| |||||||||||| |||||||||||||||| ||  || |||||||||
Db 70075 TCAAAGGATCCTCCCACCTCAGCCTCCCGAGTAGCTGGGACTATAGGTGCATGGCCACCAT 70134

Qy   585 ACCTGGCTAATTTTTATTTTTTAATTTTTTGAGACGAGTCTCAACTCTGTCACC 644
         ||||||| |||||||| |||||||  ||||||||||||||||||  ||||| ||
Db 70135 GCCCAACTAATTTTTTTTTT------TTTTTGAGATGGAGTTTCGCTCTTGTTGCC 70186

Qy   645 CAGGCTGGAGTGCAGTGGCGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAG 704
         |||||||||||||| ||| |||| ||||| ||||||||||| ||||| |||||||||| ||||
Db 70187 CAGGCTGGGGTGCAGTGGTGCAATGCGGCCACTGCACTCCAGCCTCTCGGGGTTCAAG 70246

Qy   705 TTATTCCTCCTGCCTCAGCCTCCCGAGTAGTTGGGACTACACACGCGCCACCACGCCCTAGCT 764
         | |||| | |||||||||||||||| |||||| |||| | || ||||||||||| |||||| 
Db 70247 CAATTCTCCTGCCTCAAGCCCCCAAGTAGTGGGATTACAGGCGCCACCACGCCTAGCT 70306

Qy   765 AATTTTTTGTATTTTTAGTAGAGANGGGG-TTCACCATGTTGCCAGGTTGATCTGAT 823
         ||||||||| |||||| || ||  || |  |||| ||||||||||||| || || 
Db 70307 GCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCAGGTTGTTCTCAAA 70366

Qy   824 CTCTGGACCT--TGTGATCTCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG 881
         ||||| ||||  | |||||| || ||||||| ||||||||||||||||||||||||||||||
Db 70367 CTCCTGACCTCAGGTGATCCGCCTGCCTCGGCCTCCCAAAGTGGGATTACAGGCGTG 70426

Qy   882 AGCCACCACCGCCCGGCCTT---ATTTTTAAATTTTTGTTTAAATGGAATCTCACTCT 938
         |||||||||| |||||||||   |||||||||||| | |||| |||| ||||| ||||
Db 70427 AGCCACTGTGCCTGGCCAATTTTTTTTTTTGGAGACAAAGTCTCGCT 70486

Qy   939 GTTACCCAGGCTGGAGTGCAAATGCCAAAATCTGCCTCAGCCTCCCAAGCTGGGATTAC 998
         ||||||||||| |||||| |||| ||| ||||||||||||||||||| ||||| |||||
Db 70487 CTTTCCCAGGCTGGAGTGGCAATGGCGCAACTCTCGGCTCACTGCAACCTCTGCCTTCGGG 70546

Qy   999 CTCAAGGGATTCTCCTGCCTCTGCCTCCCAAGCCACCTGCCTCCCACCA 1058
         ||||||||||||||||||||||||||||||||| |||| ||||||||||
Db 70547 TTCAAGCGGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTAGCACAGGCCATTGCCACCA 70606
```

```
Db 70547 TTCAAGCGGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTAGCACAGGCCATGTGCCACCA 70606
         |||||| || | |||||| |||||||||||||||||||| || || || ||| || |||||||||
Qy  1059 CACCCGCGCTAATTTTTGTATTTTCATTAGAGACGGGGTTTCACCATATTTGTCAGGCTGG 1118

Db 70607 CGCCCGGCGCTAATTTTTGTATTTTTAGTAGAGACGGGGGTTTCACCATGTTGGCCAGGCTGG 70666
         ||||| |||||||||||||||||| ||||||||||||||  |||||| ||||| ||||||||
Qy  1119 TCTCAAAACTCCTGACCTCAGGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC 1178

Db 70667 TCTTGAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC 70726
         || | ||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Qy  1179 AGGCGTGAGCCACCGCACCCACCAGCC-----GGCTAATTTAGATAAAAAAATAATGTAGCAATG 1234

Db 70727 AGGCATGAGCCACCACACCCACTGGGCCCTGCCCAAACTAATTAAAAAAAAATTTTTGTGAGA 70786
         |||| |||||||||| ||||||  |||    ||||||||||||  |||||||||||||  |||
Qy  1235 GGGGGTCTTGCTATGTTGTTCCCAGGCTGGTCTCGAACTCAGATGATCCTG 1294

Db 70787 CAGGATCTCACTATGTTGTTCCAGGACTGGTCTTGAACTCCTGGGCTTAAGGGATCGTCCTG 70846
         ||||| ||| ||||||||||||| ||| |||||| |||||| |||  |||||||| || |||
Qy  1295 AAATGAGCCAC 1304

Db 70847 TCTCAGCCTC 70856
```

<div></div>

```
RESULT 11
LOCUS       AX330738/c            65608 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  AX330738 Sequence 1247 from Patent WO0194629.
ACCESSION   AX330738
VERSION     AX330738.1  GI:18103716
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,G.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 1247 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT    15987 a  16979 c  16761 g  15878 t      3 others
ORIGIN

Query Match  40.8%; Score 587.8; DB 6; Length 65608;
Best Local Similarity  70.6%; Pred. No. 1.2e-165; Mismatches 347; Indels 44; Gaps 10;
Matches  940; Conservative  0;

Qy     4 TTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTTGCCAGGCTGGAGTGCAATGGCGCAATCT 63
         |||| |||||||||||||||| ||||||||||||||||||||||| |||| ||||| | |||
Db 39993 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTTGCCAGGCTGAAGTGTAATGGCACAGATCT 39935

Qy    64 CAGCTCACCGCGAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCCAG 123
         ||||||||| | ||| |||||||||| |||||||||| ||||||| |||||||||||||| 
Db 39934 CAGCTCACCAACAACTTCCGCCTCCCAGGTTCAAGCAATTCTCCTACCTCAGCCTCCCGAG 39875

Qy   124 TAGCTGGGATTACAGGCATGCGCCACCACGCCCGGCTAATTTTTGTATTTTTTTAGTAG 183
         |||||||||||||||||||| |||||||||||||||||||||||||||||||   ||| 
Db 39874 TAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTT---AGT 39819

Qy   184 AGATGGAGTTTCTCCATGTTGGTGCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCCT 243
         ||||||||||||| ||||||||| |||||||||||||||||||||||| |||||| ||||| 
Db 39818 AGACGGGGTTTCTCCATGTTGGTTCAGGCTGGTCTCGAACCCCCGACCTCAGATGATCCGC 39759

Qy   244 CCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTACAGGCGCCCGGCCATCTGG 303
         |||||||||||||||||||||||   | | ||||||||| |||||||| |||||
Db 39758 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTTACAGGAAAACATTTA 39700

Qy   304 CTAATTTTTGTGGTAGAAACAGGGTTTCACTGATGATGTGTCCCAAGCTGGTCTCGAGCTCA 363
         |||||||||||| |||||||||||||||||||| |  | | ||  |||||||  || |||||
Db 39758 CCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGTTTCACTGATGATGTGTGAGTGTTTTACAAGAAAACATTTA 39700
```

RESULT 12

```
LOCUS       AX332242              65608 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 2751 from Patent WO0194629.
ACCESSION   AX332242
VERSION     AX332242.1  GI:18122876
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 2751 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT    15987 a  16879 c  16761 g  15878 t      3 others
ORIGIN
```

```
Query Match            40.8%;  Score 587.8;  DB 6;  Length 65608;
Best Local Similarity  70.6%;  Pred. No. 1.2e-165;
Matches  940;  Conservative   0;  Mismatches  347;  Indels  44;  Gaps  10;
```

```
Qy      4 TTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGTCCCAGGCTGGAGTGCAAATCT     63
          |||||||||||||||||||||||||||| |||||||||||||||||||| ||||||
Db  39993 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTCCCAGGCTGGAGTGTCAAATGGCAGATCT  39935

Qy     64 CAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG    123
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39934 CAGCTCACCACAACTTCCGCCTCCCAGGTTCAAGCGAATTCTCACCTCACCTGCCTCAGCCTCCCGAG  39875

Qy    124 TAGCTGGGAATACAGGCATGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG    183
          |||||| ||||||||||||| ||||||||||||||| |||||||||||||||||| ||||
Db  39874 TAGCTGGGATTACAGGCATGTGCCACCACACCTGGCTAATTTTTGTATTTTTTAGTAG  39819

Qy    184 AGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGGTGATCCT    243
          |||| ||||||||||||||||| |||||||||||||||||||||||||||||| ||| |
Db  39818 AGACGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCCGC  39759

Qy    244 CCGTCTCGGCCTCCCAAAGTGCTAGGATTACAGGACTGGCCACCATGCCCGGCTCGCTGG    303
          ||||||||||||||||||||||||| |||||||||| |||||||||| ||||||||| ||
Db  39758 CCACCTTGGCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGTTTACAAGAAAACATTTA  39700

Qy    304 CTAATTTTTGTGGTGTAGAAAACAGGGGTTTCACTGATGTGCCCAAGTGGTCTCTGAGCTCA    363
          ||||||||||||||||||||||||||| || || ||||||||||| ||| ||||||||||
Db  39699 AGTATGAAAGACAGTACCTCCGGAATATGCACGGCGCCCTCCAAAACCCAGCCTTGAAGACTC  39640

Qy    364 AGCAGTCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTA----------    403
          |||||||||| |||||||||||||||||||||||||| ||
Db  39639 TTCACTGTCCTGGCTGGGGGGCCACCACTGCCTCCACCTCAAAAACTGGGCTGTGTTCTCTT  39580

Qy    404 CAGGCGTGCAGCCGGTGGCCGGTGCCTGGGCCTTTGTTATTTATTTATTTATTAAGACAGAGTTCGTTT    463
          |||||| | |||||||||||||| ||| ||||||||||||||||||||||||| |||||||||||||||
Db  39579 CCCCCTCCATCCCTGCCTTTTTTTTTTTTTTTTTTTTTAAGACACAGGTGTTT  39520

Qy    524 GATC-AAGCATCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGCACCACT    582
          |||| |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db  39519 TGTTGCCCAGGCTGGAGTGCAGTGGTGCAAACATGGTCACTGATCTCTTGG  39460

Qy    583 ACACCTGGCTAATTTTTATTTTTATTTTTGAGACAGAGTCTCAACTCGTCA    642
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39459 GCTCAAAGGATCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTATAGGTGCCACC  39400

Qy    643 CCCAGGCTGGAGTGCAGTGGCGCGAGTCCGGCGCCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCA    702
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39399 ATGCCCAACTAATTTTTTTTTTTT------------TTTTTGAGATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTTG  39348
```

```
RESULT 13

LOCUS       AX335496                6005 bp    DNA     linear   PAT 09-JAN-2002
AX335496/c
DEFINITION  Sequence 6005 from Patent WO0194629.
ACCESSION   AX335496
VERSION     AX335496.1  GI:18126215
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
            Horrigan,S., Soppet,D.R. and Weaver,Z.
  TITLE     Cancer gene determination and therapeutic screening using signature
            gene sets
  JOURNAL   Patent: WO 0194629-A 6005 13-DEC-2001;
            Avalon Pharmaceuticals (US)
FEATURES             Location/Qualifiers
     source          1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT    15987 a  16679 c  16761 g  15878 t      3 others
ORIGIN
```

```
Query Match             40.8%;  Score 587.8;  DB 6;  Length 65608;
Best Local Similarity   70.6%;
Matches  940;  Conservative    0;  Mismatches  347;  Indels   44;  Gaps   10;
```

```
Qy       4 TTTTTTTTTGAGATGGAGTTTTTTCGGCTGTTGCTCAGGCTGGAGTGCAATGGCGCAATCT 63
           |||||| ||||  ||  ||||||||  ||  ||||||||||||||||||||||||||| |||
Db   39993 TTTTGTTTTGAGATGGAGTTTT-GCTCTTGTTGCCCAGGCTGGAGTGCAGTGGCACGATCT 39935

Qy      64 CAGCTCACCGCAAACCTCCGCCTCCCCGGGTTCAAGCGAGTTCTCCTGCCTCAGCCTCCCG 123
           |||| |||| ||||||||| || |||||||||||  || ||| |||||||||||||||| |
Db   39934 CAGCTCACCACAACTTCCGCCTCCCCAGGTTCAAGCGAGTTCTCCTGCCTCAGCCTCCCGAG 39875

Qy     124 TAGCTGGGATTACAGGCACATGTGCCACCACGCGTCGGCTAATTTTTGTATTTTTTAGTAG 183
           ||||||||||||||||| |||||  |||||||| ||  |||||||||||| |||| |||||
Db   39874 TAGCTGGGATTACAGGCGTGCACCACCACGCGCCTAATTTTTGTATTTTT---AGT 39819

Qy     184 AGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATCCCT 243
           |||||||||| | ||||| ||||||||||||| |||||||| |||||| |||||| ||||
Db   39818 AGACGGGGTTTCTCACATGTTGGTCAGGCTGGTCTTGAACCCCTGACCTCAGATGATCCGC 39759

Qy     244 CCGTCTCGGCCCTCCCAAAGTGCTGAGATTACAGGACTGGCCACCATGCCCGGCCTGCCTGG 303
           || ||||||||||||||||||||||| |  |||||||||||||||||||||||||||||
Db   39758 CCACCTTGGCCCTCCCAAAGTGCT-GGGATTACAGGCGTGAGCCACCGCGCCCGGCCTCG 39700

Qy     304 CTAATTTTTGTGTAGAAAACAGGGTTTCACTGATGTTGCCCAGGCTGGTCTCGAACTC 363
           |||||||||||||| || ||| |||||||||| |  ||| ||||||||||| |||||
Db   39699 AGTATGAAAGCAGTACCTCCGGAATATGCACGGCGCCCTCCAAAACCAGCCTGTTGTAGACTC 39640

Qy     364 AGCAGTCCACCGTCCCTCAGCCTCCCAAAGTGCTGGGATTA--------------- 403
           || |||||||| || ||||||||||||||||||||||||||
Db   39639 TTCACTTGTCTGGCCCTGGGCCCACCACTCAGCTCAACAAACTGGGGTGTGTTTCTT 39580

Qy     404 CAGGCGTGCAGCCGTGCCTGGCCTTCACCATGCCTTTCAACTCTGTTCA 463
           ||||||| |||||| |||||| || |||||||   |||||||||| ||
Db   39579 CCCCCCATCCTGCGTTTTTTATTTTTTATTTTTTTGAGATGGGAGTCTCGC 39520

Qy     524 GATC-AAGCATCCTCGCCTCAGCCTCCCAAGTAGCTGGGACCAAGCACATGCCACCACT 582
           |||| ||| ||||||||||||||||||||| |||||||||||||||||||||||
Db   39519 TGTTGCCCAGGCTGGAGTGCAGTGGCAAACAATGGCACTCACTGCAGCCTTGAACTCTTGG 39460

Qy     583 ACACCTGGCTAATTTTTTTATTTTTTATTTTTTTGAGATGGGAGTCTCGCTGTCA 642
           |||||||| |||||||||||||||||||||||
Db   39459 GCTCAAAGGAATCCTCCCCACCTCAGCCTCCTGAGTAGCTGGGACTATAGGTGCATGCCACC 39400

Qy     643 CCCAGGCTGGAGTGCAGTGGCGCCAATCGTTGGCTCACTGCAACCTCTGCCTCCGGGGTTCA 702
           ||||||||||||||||||||| |||||  ||||||||||||||||||||||||||||||
Db   39399 ATGCCCAACTAATTTTTTTTTTTTTTTT------TTTTTGAGATGGAGTTTCGTCTTGTTG 39348

Qy     703 AGTTATTCTCCTGCCTCAACCCCCAGTAGCTGGGATTACAGGCGCCATCACACC 762
           |||| |||||||||||| ||  ||||||||||||||||||||  ||||||||||
Db   39347 CCCAGGCTGGAGTGCAATGGCGCGATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCA 39288

Qy     763 CTAATTTTTTTGTATTTTTAGTAGAGATGGGG-TTCACCATGTTCGCCAGGTTGATCTTG 821
           |||||||| || ||||||||||||||||||||  |||||||||||||||||||||||||
Db   39287 AGCAATTCTCCTGCCTCAACCCCCAGTAGCTGGGATTACAGGCGCCATCACACC 39228

Qy     822 ATCTCTGACCT--TGTGATCTGCCTCGCCTCCCAAAGTGCTGGGATTACAGGCG 879
           |||||||||||  || ||||  ||||||||||||||||||| ||||||||||||
Db   39227 TGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACTGATGTTGATCTTG 39168

Qy     880 TGAGCCACCACGCCCGGCCGCTT--ATTTTTTGTTTTGTTTTGTTTTGTTTTGAATCCACTC 937
           ||||||||||||||||||| ||||  || ||||||  ||||| ||| |  |||||||
Db   39167 AACTCCTGACCTCAGGAGGATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGTG 39108

Qy     938 TGTTACCCAGGCTGGAGTGCAATGGCCAAAATCTCGGCTCACTGCTCCCCG 997
           |||||||||||||||||||||  || ||  || ||||||||||||||||||
Db   39107 TGAGCCACTGTGCCCTGGCCTGCCAAATTTTTTTTTTTTTTTGGAGACAAAGTCTCGG 39048

Qy     998 GCTCAAGCGGATTCTCCTGCCTCAGCCTCCCAAGCAGCTGGGATTACAGGGCACCTGCCACC 1057
```

```
Db  38987 GTTCAAGCAGATTCTCCTGCCTCAGTCTCCCAAGTAGCTAGGACTACAGGCAATGTGCCACC 38928
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   1058 ACACCCGCGTAATTTTTTGTATTTTTCATTAGGAGGCGGGTTTCACCATATTTGTCAGGCTG 1117
          | ||||  || |||||||||||||| ||| ||||||| ||| |||||| || || || |||
Db  38927 ACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGAGGTTTCACCATGTTGGCCAGACTG 38868

Qy   1118 GTCTCAAACTCCTGACCTCAGGTGACCGACCTCGCCTTCCAAAGTGCTGGGATTA 1177
          ||||| || ||| ||||| |||| ||| ||||| |||||||||||||||||||||
Db  38867 GTCTTGAACTCCTAACCTCAGGTGATCCACCCGCCTTCCAAAGTGCTGGGATTA 38808

Qy   1178 CAGGCGTGAGCCACCACCTCACCCAGCC----GGCTAATTTAGATAAAAAAAATATAGCAAT 1233
          |||||| ||||||||||||||||||||||    |||||||||||||| |||||||||||||||
Db  38807 CAGGCGTGAGCCACCACCTCACCCAGCCTTCTGCCAACTAATTAAAAAAAAATTTTGTAGAG 38748

Qy   1234 GGGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAACTCCTGACCTCATGCAAATCCTTCC 1293
          |||  |||||| ||||||| |||||||||||||||||||||| |||| |||  ||||| ||||
Db  38747 ACAGGACTCACTATGTTGTCCAGACTGGTCTTGAACTCCTGGGCTTAAGGGATCGTTCCT 38688

Qy   1294 AAATGAGCCAC 1304
          ||||| |||||
Db  38687 GTCTCAGCCTC 38677
```

```
LOCUS       HSU62293       65608 bp    DNA     linear   PRI 26-OCT-1996
DEFINITION  Human LIM-kinase1 and alternatively spliced LIM-kinase1 (LIMK1)
            gene, complete cds.
ACCESSION   U62293
VERSION     U62293.1  GI:1432163
KEYWORDS    Williams syndrome.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 65608)
  AUTHORS   Frangiskakis,J.M., Ewart,A.K., Morris,C.A., Mervis,C.B.,
            Bertrand,J., Robinson,B.F., Klein,B.P., Ensing,G.J., Everett,L.A.,
            Green,E.D., Proschel,C., Gutowski,N.J., Noble,M., Atkinson,D.L.,
            Odelberg,S.J. and Keating,M.T.
  TITLE     LIM-kinase1 hemizygosity implicated in impaired visuospatial
            constructive cognition
  JOURNAL   Cell 86 (1), 59-69 (1996)
  MEDLINE   96291399
   PUBMED   8689688
REFERENCE   2  (bases 1 to 65608)
  AUTHORS   Frangiskakis,J.M., Odelberg,S.J., Atkinson,D.L. and Keating,M.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUN-1996) Human Genetics, Univ. of Utah, 10 N 2030 E,
            Bldg 533, Suite 2100, Salt Lake City, UT 84112, USA
FEATURES             Location/Qualifiers
     source          1..65608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /map="7q11.23"
                     /tissue_type="placenta; hippocampus"
                     /note="Stratagene catalog No. 936205; chromosome
                     7-specific flow-sorted cosmid library from Lawrence
                     Livermore National Laboratories"
     gene            1..38534
                     /gene="LIMK1"
     mRNA            join(<1..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37166..38524)
                     /gene="LIMK1"
                     /product="LIM-kinase1"
                     /note="Williams syndrome region; Kiz-1; protein kinase
                     with two LIM domains"
     CDS             join(47..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37166..37328)
                     /gene="LIMK1"
                     /codon_start=1
                     /product="LIM-kinase1"
                     /protein_id="AAB17545.1"
                     /db_xref="GI:1432164"
                     /note="Williams syndrome region; Kiz-1; protein kinase
                     with two LIM domains"
                     /translation="MRLTLLGCCTWREERMGEEGSELPVCASCGQRIYDGQYLQALNAD
                     WHADGERCCDSASLSHQYYEKDGQLFCKKDYWARYGESCHCSQITKGLVWVAGEL
                     KYHPECFICLICGTFIGDGDTYTLVEHSKLYCGHCYQTVTPVIEQILPDSPGSHLP
                     HTVLVSIPASSHGRGLSVSIDPPHGPPGCSTEHSHTVRVQGVDGPCMSPDVKNSIH
                     VGDRLIEINGTPIRNVPLDEIDLLIQETSRLLQITLEHDPHDTLGHGLGPETSPLSSP
                     AYTPSGEAGSSAROKPVLRSCSIDRSPGAGSLSGPSASQRKDLGRSESLRVVCRPHRIF
                     RPSDLIHGEVLGKGCFGQVAIKVTHRETGEVMVMKELIRFDEETQRTFLKEVKVMRCLE
                     HPNVLKFIGVLYKDRKLNFITEYIKGGTLRGIIKSMDSQYPWSQRVSFAKDIASGMAY
                     LHSMNIIHRDLNSHNCLVRENKNVVADFGLARLMVDEKTQPEGLRSLKKPDRKKRYT
                     VVGNPYWMAPEMINGRSYDEKVDVFSFGIVLCEIIGRVNADPDYLPRTMDFGLNVRGF
                     LDRYCPPNCPPSFYPITVRCCDLDPEKRPSFVKLEHWLETLRMHLAGHLPLGPQLEQL
                     DRGFWETYRGESGLPAHPEVPD"
     polyA_site      join(89..101,1789..1885,12663..12801,13121..13230,
                     15072..15278,21920..22025,22122..22288,23055..23238,
                     23915..24001,24949..25080,27692..27751,27977..28042,
                     31846..32002,36610..36665,36919..37076,37166..37328)
     misc_feature    62161
                     /gene="LIMK1"
                     /note="K2049 deletion breakpoint"
     polyA_site      38524..38529
                     /gene="LIMK1"
     polyA_site      38529..38534
                     /gene="LIMK1"
BASE COUNT    15987 a   16679 c   16761 g   15878 t      3 others
ORIGIN
```

```
Query Match        40.8%;  Score 587.8;  DB 9;  Length 65608;
Best Local Similarity   70.6%;  Pred. No. 1.2e-165;
Matches  940;  Conservative   0;  Mismatches  347;  Indels  44;  Gaps  10;

Qy      4 TTTTTTTTTGAGATGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAATCT 63
          |||| ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db  39993 TTTTTGTTTGAGATGGAGTTTT--GCTCTTGTTGCCCAGGCTGGAGTGCACGATCT 39935

Qy     64 CAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAG 123
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39934 CAGCTCACCACAACCTCCGCCTCCCAGGTTCAAGCAATTCTCCTCAGCCTCCCGAG 39875

Qy    124 TAGCTGGGATTACAGGCACATGTGCACCACCACGCCCGGCTAATTTTTGTATTTTTAGTAG 183
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||| |||
Db  39874 TAGCTGGGATTACAGGCATGTGTCACCACCACTGGCTAATTTTTTGTATTTT---AGT 39819

Qy    184 AGATGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGTGATGATCCCT 243
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39818 AGACGGGGTTTCTCCATGTTGGTCAGGCTGTTGGTCTTGAAACCCCGACCTCAGATGATCCGC 39759
```

```
Qy    244 CCGTCTCGGCCCTCCCAAAGTGTCTAGATACAGGACTGGGCCACCAATGCCCGGCTCTGCCTGG 303
          || || |||||||||||||||||||||||||||| || |||||||||| |||||||| |||||
Db  39758 CCAACCTTGGCCTCCCCAAAGTGCT--GGGATTACAGGTGTGAGTTTTACAAGAAAACATTTA 39700

Qy    304 CTAATTTTTTGTTGGTAGAAACAGGGTTGTCACGGCTGGTCTCCTGAGCTCA 363
          | ||||||||||||||||||||||||||||||||| |||||| |||||||||
Db  39699 AGTATGAAAGACAGTACCTTCCGAGAATATGCACGGCCGGTCTCCTGAGCTCA 363

Qy    364 AGCAGTCCACCTGCCTCCAGCCGTCCCAAAGGTCTGGGATTA------------- 403
          |-|| |||||||||||||||||||||||||||||||||||||
Db  39639 TTCACTTGTCTCTGGGCTGGGCCACCACCTCCACCTAAAACTGGGCGTGTGTTCTCTT 39580

Qy    404 CAGGGCGTGCAGCCGCCGTGCCTGGCCTTTTTATTTTTATTTTTTTAAGACACAGGTGTCCCAC 463
          || |||| |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  39579 CCCCCCCATCCTGCCCCTTTTTTATTTTTATTTTTTTTTGAGACAGAGTCTTTGTTT 39520

Qy    464 TCTTACCCAGGATGAAGTGCAGTGGCGTGATCACAGCTCACTGCAGCCTTCAACTCCTGA 523
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39519 TGTTGCCCAGGCTGGAGTGCAGTGGCTCACTGCAACAATGGCTCACTGCAGCCTCTGACTCCTTGG 39460

Qy    524 GATC-AAGCAATCCTCCCAGCCTCCCAAGTACTAGCTGGGACCAAAGACATGCACCACT 582
          |||| |||||||||||||||||||||||||||||||||||||||||||| ||||||
Db  39459 GCTCAAAGGATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTATAGGTGCATGCCACC 39400

Qy    583 ACACCTGGCTAATTTTTTTTTTTTTATTTTTAGTAGAGACAGAGTCTCAACTCTGTCA 642
          || ||||||||||||||||||||||||||| |||||||||||||||||||||| |||
Db  39399 ATGCCCAACTAATTTTTTTTTT--TTTTTGAGAATGGAGTTTCGCTCTTGTTG 39348

Qy    703 AGTTATTCTCTCGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCAGCCCTAG 762
          |||||||| |||||||||||||||||||||| |||||
Db  39347 CCCAGGCTGGGGTGCAAATGGTGCGATCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCA 39288

Qy    763 CTAATTTTTTTGTAGTTTGTATTTTTAGTAGAGATGGGG--TTCACCATGGTTGGATCTTG 821
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db  39287 AGCAATTCTCCTGCCTCAACCCCCCAAGTAGCTGGGATTACAGGCATGTGCCATCACAACC 39228

Qy    822 ATCTCTGGACCT--TGTGATCTGGCCGCTCCCAAAGTGCTGGGATTACAGGCG 879
          |||| |||| ||||||||||||||||||||||||||||||||||||||||||
Db  39227 TGGCTAATTTTGTATTGTAGTAGAGATGGGGTTTCACTATGGTTCCAGGCTGGTCTCCA 39168

Qy    880 TGAGCCACCACGCCCGGCCTT--ATTTTTAATTTTTGTTGTTTGAAATGGAATCTCACTC 937
          |||||||||||||||||||| ||||||||||||||||||||||||||
Db  39167 AACTCCTGACCTCAGGATGATCCTCCTCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGTG 39108

Qy    938 TGTTACCCAGGCTGGAGTGCAAATGGCCAAATCTCGGCTCACTGCAACCTCCGG 997
          |||| |||||||||||||||||||||||||||||||||||||||||||||
Db  39107 TGAGCCACTGTGCCCGGCCCAATTTTTTTTTTTTTTTTTGGAGCAAAAGTCTCGC 39048

Qy    998 GCTCAAGCGATTCTCTCCTGTTCTCCAGCCTCCCAAGTACGGGCACCTGCCACC 1057
          |||||||||| ||||||||||||||||||||| |||||||||||||||||
Db  39047 TCTTTCCCAGGCTGGAGTGGAAATGGCACCACCTCTCGCCTTCCGG 38988

Qy   1058 ACACCCGGCTAATTTTTTGTTAATTTCATTAGAGGCGGGGGTTTCACCATATTTGTCAGGCTG 1117
          || |||||||||||||||||||||||||||||||||||||||||||
Db  38987 ACGCCCGGCTAATTTTTGTATTTTTAGTAGAGACGAGTTTCACCATGTTGGCCAGGCTG 38928

Qy   1118 GTCTCAAACTCCTGACCTGAGGTGACCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTA 1177
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  38927 ACGCCCGGGCTAATTTTTGTATTTTTAGTAGAGACGGGTTTCACCATGTTGGCCAGACTG 38868

Qy   1178 CAGGGCGTGAGCCACCTCACCCAGCC----GGCTAATTTAGATAAAAAAATATGTAGCAAT 1233
          |||||||||||||||||||||||| |||
Db  38867 GTCTTGAACTCCTAACTTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTA 38808

Qy   1234 GGGGGGTGTCTGCTAGTGTTGCCCAGGCTGGTCTCAAACTTCTGGCTTCATGCAATCCTTCC 1293
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  38807 CAGGCATGAGCCACCACCACACCTGGCCTCTGCCAACTAATTAAAAAAAATTTTTGTAGAG 38748

Db  38747 ACAGGATCTCACTATGTTGTCCAGGACTGGTCTTGAACTCCTGGGCTTCAAAGGGATTCGTCCT 38688
```

```
Qy   1294 AAATGAGCCAC 1304
          |||| |||||
Db  38687 GTCTCAGCCTC 38677
```

```
                     /translation="LLNLAGLVPGVGVAPGVGVAPGVGVAPGVGVLAPGVGVAPGVGVA
                     PGVGVAPGIGPGGVAAAAKSAAKVAAAKAQLRAAAGLGAGIPGLGVGVPGLGVGAGV
                     PGLGVGAGVPGFGVAVPGALAAAKAAKYCAAVPGVLGGLGALGVGIPGGVGGLGVPGGLGGIPPAAAAKAAKYGAAGLGG
                     AAAKAAAKAAQPGLVGGAAPGLVGGAAGLGGLGVGGLGVPGVGGLGGIPPAAAAKAAKYGAAGLGG
                     VLGGAGQPFLGVAARPGFGLSPIFPGGACLGKACGRKRK"
repeat_region        1370. .1798
                     /note="91% similarity"
                     /rpt_family="ALU"
repeat_region        2472. .2596
                     /note="70% similarity"
                     /rpt_family="MIR"
repeat_region        complement(2717. .2985)
                     /note="91% similarity"
                     /rpt_family="ALU"
repeat_region        3027. .3139
                     /note="81% similarity"
                     /rpt_family="ALU"
repeat_region        3964. .4251
                     /note="83% similarity"
                     /rpt_family="ALU"
repeat_region        4282. .4737
                     /note="92% similarity"
                     /rpt_family="ALU"
repeat_region        5094. .5370
                     /note="81% similarity"
                     /rpt_family="ALU"
repeat_region        5479. .5766
                     /note="90% similarity"
                     /rpt_family="ALU"
repeat_region        6424. .6708
                     /note="91% similarity"
                     /rpt_family="ALU"
repeat_region        7150. .7912
                     /note="90% similarity"
                     /rpt_family="ALU"
repeat_region        complement(7408. .7692)
                     /note="93% similarity"
                     /rpt_family="ALU"
repeat_region        7698. .8085
                     /note="92% similarity"
                     /rpt_family="ALU"
repeat_region        9786. .9791
                     /rpt_family="ALU"
polyA_signal         10016. .10021
                     /gene="ELN"
polyA_signal         10306. .10445
                     /gene="ELN"
repeat_region        complement(11671. .11782)
                     /note="78% similarity"
                     /rpt_family="ALU"
repeat_region        11829. .12107
                     /note="79% similarity"
                     /rpt_family="ALU"
repeat_region        complement(12129. .12193)
                     /note="92% similarity"
                     /rpt_family="ALU"
repeat_region        12545. .12817
                     /note="88% similarity"
                     /rpt_family="ALU"
repeat_region        complement(12971. .13260)
                     /note="88% similarity"
                     /rpt_family="ALU"
repeat_region        13645. .13747
                     /note="87% similarity"
                     /rpt_family="ALU"
repeat_region        complement(13841. .14132)
                     /note="87% similarity"
                     /rpt_family="ALU"
repeat_region        14274. .14554
                     /note="88% similarity"
                     /rpt_family="ALU"
repeat_region        14656. .14736
```

```
repeat_region        1483. .14929
                     /rpt_family="MLT1"
repeat_region        14933. .15566
                     /note="85% similarity"
                     /rpt_family="ALU"
repeat_region        15590. .15878
                     /note="71% similarity"
                     /rpt_family="L1"
repeat_region        16173. .16458
                     /note="65% similarity"
                     /rpt_family="ALU"
repeat_region        16791. .16995
                     /note="92% similarity"
                     /rpt_family="ALU"
repeat_region        17035. .17809
                     /note="69% similarity"
                     /rpt_family="L1"
repeat_region        17922. .18035
                     /note="90% similarity"
                     /rpt_family="ALU"
repeat_region        18261. .18538
                     /note="82% similarity"
                     /rpt_family="MLT1"
repeat_region        complement(18545. .18714)
                     /note="88% similarity"
                     /rpt_family="ALU"
repeat_region        19070. .19194
                     /note="82% similarity"
                     /rpt_family="ALU"
repeat_region        19696. .19989
                     /note="76% similarity"
                     /rpt_family="ALU"
repeat_region        20022. .20317
                     /note="88% similarity"
                     /rpt_family="ALU"
repeat_region        20351. .20626
                     /note="79% similarity"
                     /rpt_family="ALU"
repeat_region        complement(20823. .21424)
                     /note="88% similarity"
                     /rpt_family="ALU"
repeat_region        complement(22002. .22167)
                     /note="91% similarity"
                     /rpt_family="ALU"
repeat_region        22188. .22482
                     /note="79% similarity"
                     /rpt_family="ALU"
gene                 23511. .62185
                     /note="89% similarity"
                     /rpt_family="ALU"
mRNA                 join(23511. .23730,25433. .25529,36304. .36442,36762. .36871,
                     38711. .38917,45557. .45562,45759. .45925,46692. .46875,
                     47552. .47638,48586. .48717,51327. .51386,51612. .51677,
                     55481. .55637,60244. .60299,60553. .60710,60800. .62185)
                     /gene="LIMK1"
exon                 /product="LIM-kinase"
                     /note="determined using HUMLIMKA, Mizuno et al., Oncogene
                     9,1605-1612,1994"
                     9,1605-1612,1994"
                     23511. .23730
```

```
Query Match          40.8%;   Score 587.8;   DB 9;   Length 67046;
Best Local Similarity   70.6%;   Pred. No. 1.2e-165;
Matches 940; Conservative 0; Mismatches 347; Indels 44; Gaps 10;

Qy          4 TTTTTTTTTGAAGATGGAGTTTCGCTCCTTGTTGCCCAGGCTGGAGTGCAATGGCGCAATCT 63
              |||| ||||||||||||||||||| || ||||||||||||| ||| ||||| |||||||||
Db      63627 TTTTGTTTGAGATGGAGTTTT-GCTCTTGTTGTTGCCCAGGCTGGAGTGTAATGGCACGAATCT 63569

Qy         64 CAGCTCACCGCGAACCTCCGCCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCCAG 123
              ||||||||||||||| ||||||| ||||||||||||||||||| |||||||||||||||||
Db      63568 CAGCTCCACCACAACTTCCGCCTCCGCCTCCCCAGGTTCAAGCAATTCTCCTACCTCAGCCTCCCGAG 63509
```

```
Qy    124  TAGCTGGGATTACAGGCATGTGCACCACGCGCTCGGCGTTAATTTGTATTTTTTAGTAG  183
           ||||||||||||||||||||||||||||| ||||| | |||||||| |||   ||||
Db  63508  TAGCTGGGATTACAGGCATGTGCACCACACTGGCTAATTTTTTTATTTT---AGT  63453

Qy    184  AGATGGAGTTTCTCATGTTGGTCAGGCTGGTCTCGAACTCCGACCTCAGATGATCCCT  243
           | || || |||||||| ||||||||||||||||| | |||||||||||||||||||
Db  63452  AGACGGGGTTTCTCCATGTTGGTCAGGTTGGTCTTGAACCCCGACCTCAGATGATCCGC  63393

Qy    244  CCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACTGGCCACCATGCCCGGCCTCCCTG  303
           ||| || ||||||||||||||||| || | |||| |||||||||||||||||||||
Db  63392  CCACCTTGGGCTCCCAAAGTGCCT-GGGATTACAGGCGTGAGTTTTACAGAGAAACATTTA  63334

Qy    304  CTAATTTTGTGTAGAAACAGGTTTCACTGATGTGCCCAGCCTGGTCTCCTGAGCTCA  363
           |||||||||| | ||| |||||||||||||||| ||||||||||||||| ||||
Db  63392  AGTATGAAAGCAGTACCTCCGGAATATGCACGGCGCCCTCCAAACCAGCCTCGAAGACTC  63274

Qy    364  AGCACAGTCCACCTGCCTCAGCGTCCCCAAAGTGCTGGGATTA----------------  403
           |||| ||||||||||||| ||||||| | ||||||||||||
Db  63333  TTCACTTGTCTCTGGCTGGGGCCACCACTCAGCTCACCTAAAAACTGGGCTGGGGTCTCTT  63214

Qy    404  CAGGCGGTTGCAGCCGTGCGGCCCTTTTATTTATTTTAATTTTTTGAGACAGAGTCTCA  463
           ||||| || |||||||||| ||||| || |||||| |||||||||||||||||||
Db  63213  CCCCCCTCCATCCCCTGCCTTTTTTTTTTTTTTTGAGATGGGAGTCTTGTTCCCAC  63154

Qy    464  TCTTACCAGGATGAAGTGCAGTGGTGCGATCACAGCTCACCGCCTTCAACTCCTGA  523
           ||||||||||| ||||||| |||||||||||||||||||||||||||||||||||||
Db  63153  TGTTGCCCAGGCTGGAGTGCAGTGGTGCAAACATGGCGTCACTGCAGCCTTGACCTCTTGG  63094

Qy    524  GATC-AAGCAATCCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACCAAAGACATGCACCACT  582
           |||| ||| | |||||||||||||||||| |||||||||| || || || |||||||
Db  63093  GCTCAAAGGAATCCCACCTCACCCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGCATGCCACC  63034

Qy    583  ACAGCCTGGCCTAAATTTTAATTTTTTTATTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCA  642
           |||| ||||||| |||||| |||||| ||||||| ||||||||
Db  63033  ATGCCCAACTAAATTTTTTTTT-------TTTTTGAGATGGGAGTTTCGTCTTGTTG  62982

Qy    643  CCCAGGCTGGAGTGCAGTGGCGCGAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCA  702
           |||||||||||||||||||| | |||||||||||||||||||| || | ||||| || |
Db  62981  CCCAGGCTGGGGTGCAAATGGTGCGATCCTCGGCTCACTGCAACCTCTGCCTCCCGGGTTCA  62922

Qy    703  AGTTATTATTCTCCTGCCCAGCCTCCTGAGTAGCTGGGATTACAGGCGCCCACCACGCCTAG  762
           | ||||| ||||||||||||||||||||| || ||| |||||||||| ||| ||||
Db  62801  AACTCCTGACCTCAGGAGATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGTG  62742

Qy    763  CTAATTTTTTTGTATTTTTTAGTAGAGAATGGGG-TTCACCAGTTTCCGCCAGGTTGGTCTTG  821
           ||||||||||| | ||||| |||||||| ||||| ||||||||| ||||||||||| |||
Db  62861  TGGCTAATTTTTGTATTGTAGTAGAGATGGGGTTTCACTATGTCAGGCTGGTCTGTTCTCA  62802

Qy    822  ATCTCTGGACCT--TGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCG  879
           || ||||| ||  | |||| |||||||||||||||||||||||||||||||||||||||
Db  62801  AACTCCTGACCTCAGGAGATCCTCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGTG  62742

Qy    880  TGAGCCACCACGCCCGGCCTT--ATTTTTAATTTTTTTGTTGGTTGAAATGGAATCTCACTC  937
           ||| |||||| |||||||||   |||||||||||||| ||||||||||||||||
Db  62741  TGAGCCACTGTGCCTGGCCTGCCAATTTTTTTTTTTTTTTGAGACAAAGTCTCGC  62682

Qy    938  TGTTACCCAGGCTGGAGTGCAAATGGCCAAAATCTCGGCTCACTGCAACCTCCGCCTCCCGG  997
           |||||||||||||||| ||| ||||||||||||||||||||||||||||||||||||| |||
Db  62681  TCTTTCCCAGGCTGGAATGGCCACCACCTTGCCTTTCG  62622

Qy    998  GCTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGCACCTGGGATTACGGGCACCTGCCACC  1057
           ||||||||||||||||||||||||||||||||| || ||| |||| |||| |||||||
Db  62621  GTTCAAGGGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTAGGGCATGTGCCAC  62562

Qy   1058  ACACCCCGGCTAATTTTTGTATTTTTCATTAGAGGGGGGGTTTCACCATATTTTGTCAGGCTG  1117
           ||||||||||| |||||||||| |||| ||||||||||||||||||||||||||||||||
Db  62561  ACGCCCGGCTAATTTTTGTATTTTTAGTAGAGACCATGTTGGCCAGACTG  62502

Qy   1118  GTCTCAAACTCCTGACCTCAGGTGACCCACCTGCCTCAGCCTTCCAAAGTGCTGGGATTA  1177
           |||||||||||||| |||||||||||||||||||||| ||||||||||||||||||||
Db  62501  GTCTTGAACTCCTCCAACTTCAGGTGATCACCGCGCCTCCCAAAGTGCTGGGATTA  62442

Qy   1178  CAGGCGTGAGCCACCTCACCCAGCC----GGCTAATTTAGATAAAAAATATGTTAGCAAT  1233
           ||||| ||||||||| ||||| ||    ||||||||||| || ||||||||||||||
Db  62441  CAGGCATGAGCCACCACACCTGGCGCCTCTGCCCAACTGAATTAAAAAAAAATTTTGTAGAG  62382

Qy   1234  GGGGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCAAACTTCTGGCTTCATGCAAATCCTTCC  1293
           |||||| |||||||||||| | ||||| |||||| ||||| |||| || || ||| ||||
Db  62381  ACAGGATCTCACTATGTTGTTCCAGGACTGGTCTTGAACTCCTGGGCCTTAAGGGATCGTCCT  62322

Qy   1294  AAATGAGCCAC  1304
           ||| ||||| |
Db  62321  GTCTCAGCCTC  62311
```

Search completed: February 1, 2003, 22:45:59
Job time : 4350 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: February 1, 2003, 21:27:54 ; Search time 2217 seconds
                 (without alignments)
                 10534.006 Million cell updates/sec

Title: US-09-964-678A-1
Perfect score: 1442
Sequence: 1 tttttttttgagatggag.........ttaacaaagctttagagca 1442

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   EST:*
                 1: em_estba:*
                 2: em_esthum:*
                 3: em_estin:*
                 4: em_estnu:*
                 5: em_estov:*
                 6: em_estpl:*
                 7: em_estro:*
                 8: em_htc:*
                 9: gb_est1:*
                 10: gb_est2:*
                 11: gb_htc:*
                 12: gb_est3:*
                 13: gb_est4:*
                 14: gb_est5:*
                 15: em_estfun:*
                 16: em_estom:*
                 17: gb_gss:*
                 18: em_gss_hum:*
                 19: em_gss_inv:*
                 20: em_gss_pln:*
                 21: em_gss_vrt:*
                 22: em_gss_fun:*
                 23: em_gss_mam:*
                 24: em_gss_mus:*
                 25: em_gss_other:*
                 26: em_gss_pro:*
                 27: em_gss_rod:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	446.8	31.0	549	14	BM996475	BM996475 UI-H-DT0-
2	396.2	27.5	1842	11	BC011119	BC011119 Homo sapi
3	391.4	27.1	877	17	AQ739838	AQ739838 HS_5505_A
4	383.6	26.6	1345	10	AV762220	AV762220
5	379	26.3	910	14	BQ722917	BQ722917 AGENCOURT
6	378.4	26.2	767	10	AV700498	AV700498 AGENCOURT
c 7	373.8	25.9	658	17	AQ393450	AQ393450 CITBI-E1-
c 8	373	25.9	911	17	AQ746594	AQ746594 HS_2278_A
c 9	370	25.7	637	14	BQ438614	BQ438614 AGENCOURT
c 10	368.8	25.6	970	13	BM468547	BM468547 AGENCOURT
c 11	367.6	25.5	5556	17	AQ839814	AQ839814 260L13-C5
c 11	367.4	25.5	4087	11	BC024593	BC024593 Homo sapi
c 12	366.6	25.4	839	10	AV700988	AV700988 UI-E-E00-
c 13	366.6	25.4	671	14	BM722145	BM722145 UI-H-DH0-
c 14	364	25.2	611	11	BM996878	BM996878 UI-H-DH0-
c 15	361.4	25.1	722	14	AV761207	AV761207
c 16	360.8	25.0	930	14	AG150197	AG150197 Pan trogl
c 17	360	25.0	667	17	BQ711438	BQ711438 AGENCOURT
c 18	359.8	25.0	870	14	AG186062	AG186062 Pan trogl
c 19	359.4	24.9	687	17	BF346320	BF346320 602018525
c 20	358.8	24.9	679	12	AQ583981	AQ583981 RPCI-11-4
c 21	357.8	24.8	722	14	BQ959903	BQ959903 AGENCOURT
c 22	357.4	24.8	946	14	BC005990	BC005990 Homo sapi
c 23	357	24.8	1259	11	BQ276208	BQ276208 Homo sapi
c 24	356.6	24.7	771	14	BC017826	BC017826 Homo sapi
c 25	356	24.7	1808	11	AQ878530	AQ878530 UI-H-FH0-
c 26	355.4	24.6	740	11	BC019333	BC019333 Homo sapi
c 27	355	24.6	840	11	AQ318783	AQ318783 RPCI11-92
c 28	353.4	24.5	689	17	AL601019	AL601019 DKFZp313F
c 29	353.2	24.5	723	9	AV700545	AV700545 UI-H-DP0-
c 30	350.6	24.3	737	10	BQ000069	BQ000069 UI-H-DP0-
c 31	350	24.3	766	14	BC024155	BC024155 Homo sapi
c 32	349	24.2	626	9	AL696611	AL696611 DKFZp686K
c 33	347.8	24.1	1875	11	BC009270	BC009270 Homo sapi
c 34	347	24.1	599	14	BQ921470	BQ921470 AGENCOURT
c 35	346.6	24.0	753	9	AL691690	AL691690 DKFZp313B
c 36	346.2	24.0	1771	11	BC018833	BC018833 Homo sapi
c 37	345.6	24.0	749	10	AV699709	AV699709 UI-E-E00-
c 38	345.2	23.9	792	13	BI225234	BI225234 602499941
c 39	344.8	23.9	758	12	BG546184	BG546184 602573405
c 40	343.6	23.8	758	12	BG563032	BG563032 602581816
c 41	343.6	23.8	617	13	BI861844	BI861844 603388872
c 42	343	23.8	714	12	BG289947	BG289947 603183363
c 44	341.8	23.7	714	12	BG289947	BG289947 603183363
c 44	341.2	23.7	714	12		
c 45	341	23.6	634	17	AQ543763	AQ543763 RPCI-11-3

ALIGNMENTS

RESULT 1
BM996475

```
LOCUS       BM996475               549 bp    mRNA    linear   EST 17-JUN-2002
DEFINITION  UI-H-DT0-avl-i-16-0-UI.s1 NCI_CGAP_DT0 Homo sapiens cDNA clone
            IMAGE:5881047 3', mRNA sequence.
ACCESSION   BM996475
VERSION     BM996475.1  GI:19721376
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 549)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNl at: http://image.llnl.gov
            The following repetitive elements were found in this cDNA
            sequence: 11-305, >ALU (matched compliment) 243-535, >SVA#Other
            compliment) 221-440, >ALU (matched
            Seq primer: M13 FORWARD
```

RESULT 2
BC011119

```
FEATURES             Location/Qualifiers
     source          1..549
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5881047"
                     /clone_lib="NCI_CGAP_DT0"
                     /tissue_type="Metastatic Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     POLYA=Yes.
                     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
                     modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_DT0 is a cDNA library containing the following
                     tissue(s): Metastatic Chondrosarcoma in Lung. The library
                     was constructed according to Bonaldo, Lennon and Soares,
                     Genome Research, 6:791-806, 1996. First strand cDNA
                     synthesis was primed with an oligo-dT primer containing a
                     Not I site. Double stranded cDNA was ligated to an EcoR I
                     adaptor, digested with Not I, and cloned directionally
                     into pT7T3-Pac vector. The oligonucleotide used to prime
                     the synthesis of first-strand cDNA contains a library tag
                     sequence that is located between the Not I site and the
                     (dT)18 tail. The sequence tag for this library is
                     AACTGTTCGG.
                     TAG_LIB=UI-H-DT0
                     TAG_TISSUE=lung metastatic chondrosarcoma
                     TAG_SEQ=AACTGTTCGG"
BASE COUNT      104 a    154 c    127 g    164 t
ORIGIN

Query Match          31.0%;     Score 446.8;   DB 14;   Length 549;
Best Local Similarity 95.4%;    Pred. No. 1.1e-72;
Matches 515; Conservative  0; Mismatches 17;  Indels 8;  Gaps 5;

Qy     1 TTTTTTTTTTTTGAGATGGAGAGTTTTCGGTCTCTTGTTCGGTCTGAGTCGAGTGCAATGGCGCAA 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     8 TTTTTTTTTTTTGAGATGGAGAGTTTTCGGTCTCTTGTTCGGTCTGAGTCGAGTGCAATGGCGCAA 67

Qy    61 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGAGTTCTCTGCCTCAGCCTCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    68 TCTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGAGTTCTCTGCCTCAGCCTCC 127

Qy   121 CAGTAGCTGGGATTACAGGCGCATGTGCCACCACGCTCGGCTAATTTTGTATTTTTTAG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   128 TAGTAGCTGGGATTACAGGCGCATGTGCCACCACGCTCGGCTAATTTTGTATTTTTTAG 187

Qy   181 TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   188 TAGAGATGGGAGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATC 247

Qy   241 CCTCCGTCTCCGGCCCTCCCAAAGTGGT--AGATACAGGACGTGGCCACCACCATGCCCGG-CTC 296
           |||||||||||| ||||||||||||||   ||||| |||| ||| |||||||||||||||| ||
Db   248 CTCCCGTCTCGGGCCCTCCCAAAGTGCCTGAGATTACAGGCGTGAGCCACCATGCCCGGCCTG 307

Qy   297 TGCCTGGGCTAATTTTGTGTGTAGAAAACAGGGTTCACTGATG-TGCCCAACGTGGTGTCTCC 355
           ||||||||||||||||| ||||||| |||||||||||||||| |||||| |||||| |||||
Db   308 TGCCTGGGCTAATTTTTGTGTAGAAACAGGGTTCACTGATGTTGCCCAAGGTGGTCTCC 367

Qy   356 TGAGCTCAAGCAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGT--GCA 413
           ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||| |||
Db   368 TGAGCTCAAGCAGCAGTCCACCTGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC 427

Qy   414 GCCGTGCCTGGGCCTTTTATTTATTTTTTTAAGACACAGGGTGTCCCACTCTTACCCAG 473
           |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db   428 ACCGTGCCTGGGCCTTTTATTTATTTTTTTAAGACACAGGTGTCCCACTCTTACCCAG 487

Qy   474 GATGAAGTGCCAGTGGTGTGTGATCACAGACCCTTCAAACT-CCTGAGATCAAGCA 532
           |||||||||||||||||||| |||||||| |||| | |||| |||||||||||||||
Db   488 GATGAAGTGCCAGTGGTGTGATCACAGCTCACTGCAACCTCCTGAGATCAAGCA 547
```

```
LOCUS       BC011119            1842 bp    mRNA        linear   HTC 30-JUL-2001
DEFINITION  Homo sapiens, clone IMAGE:3047997, mRNA.
ACCESSION   BC011119
VERSION     BC011119.1  GI:15029795
KEYWORDS    HTC.
SOURCE      Homo sapiens
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1842)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Baylor College of Medicine Human Genome
            Sequencing Center
            Center code: BCM-HGSC
            Web site: http://www.hgsc.bcm.tmc.edu/cdna/
            Contact: villalon@bcm.tmc.edu.
            Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
            A.M., Holloway, M., Telford, B, Hodgson, A., Bouck, J., Yu, W.,
            Muzny,D.M., Gibbs,R.A.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 19 Row: o Column: 1
            This clone has the following problem: Incomplete processing.
FEATURES             Location/Qualifiers
     source          1..1842
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3047997"
                     /tissue_type="cervix, carcinoma"
                     /clone_lib="NIH_MGC_12"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
BASE COUNT      371 a    522 c    475 g    474 t
ORIGIN

Query Match          27.5%;     Score 396.2;   DB 11;   Length 1842;
Best Local Similarity 70.3%;    Pred. No. 1.5e-63;
Matches 746; Conservative  0; Mismatches 263;  Indels 52;  Gaps 14;

Qy     2 TTTTTTTTTTGAGATGGAGAGTTTTCGGTCTCTTGTTCGGTCTGAGTCGAGTGCAATGGCGCAAT 61
           ||||| || |||||||||||||||||||||||||||||||||||||||||||||| ||  ||
Db   570 TTTTTATTGTTTTGAGATGGAGAGTTTTCGGTCTCTTGTTCGGTCTGAGTCGAGTGCGGTGAT 629

Qy    62 CTCAGCTCACCGCAACCTCCGCCTCCCGGGTTCAAGCGGATTCTCCTGCCTCAGCCTCCC 121
           ||| ||| | ||||| ||||||| |||||||||||    ||||||||||||||||||||
Db   630 CTCGGCTTACTGCAAGCTCCGCCCTCCCGGGTTCACACCTTTCTCCTGCCTCAGCCTCCA 689

Qy   122 AGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTGTATTTTTTTAGT 181
           |||||||||||  || ||| |||||||||||||||||||||||||  |||| ||||||
Db   690 AGTAGCTGGGACTACGTGCCCGCGCCACCACGCCCGGCTAATTTT-TTTCTATTTTTAGT 748

Qy   182 AGAGATGGAGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCCGACCTCAGATGATCC 241
           ||||| ||| ||||| |||||| ||||||| |||||||||   ||| ||||   ||| ||
Db   749 AGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATGTCCTGACCT--TGTGATTC 806

Qy   242 CTCCGTCTCGGCCTCCCAAAGTGCTAGATACAGGACGTGGCCACCATGCCCGGCTCTGCCT 301
           | || |||||||||||||||||||| |||| ||||| || |||||| ||||||| |||
Db   807 ACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGAGTGAGCCACCACGCCCGGCCTTA 866

Qy   302 GGCTAATTTTGTGTGTAGAAACAGGGTTTCACTGATGTGCCCAAGCTG-GTCTCCTCAGC 360
           |||| ||||| |||||  | |  |||  ||||||||||||||||||||  |||||||||
Db   807 ...
```

RESULT 3

```
LOCUS       AQ739838/c               877 bp    DNA    linear   GSS 16-JUL-1999
DEFINITION  HS_5505_A1_C09_T7A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=1081 Col=17 Row=E, DNA sequence.
ACCESSION   AQ739838
VERSION     AQ739838.1  GI:5517360
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 877)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
  COMMENT   Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
            or from Resear h Genetics (info@resgen.com). BAC end Web Server:
            http://www.htsc.washington.edu
            Plate: 1081 row: E column: 17
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 877.
FEATURES             Location/Qualifiers
     source          1..877
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=1081 Col=17 Row=E"
                     /clone_lib="RPCI-11 Human Male BAC Library"
                     /sex="male"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     Male blood DNA was isolated from one randomly chosen donor
                     and partially digested with a combination of EcoRI and
                     EcoRI Methylase. Size selected DNA was cloned into the
                     pBACe3.6 vector at EcoRI sites"
BASE COUNT      251 a    181 c    242 g    203 t
ORIGIN
```

```
Query Match       27.1%;  Score 391.4;  DB 17;  Length 877;
Best Local Similarity   79.4%;  Pred. No. 1.5e-62;
Matches   514;  Conservative   0;  Mismatches   126;  Indels   7;  Gaps   4;
```

```
Qy    568 AAGACATGCACCACTACACTGCTAATTTTTATTTTTATTTTTTGAGACAGA 627
          ||| ||  ||     | || ||||||||| || |||| ||||| ||||||||
Db    680 AACACCTCCCACGTAATGAAATGTCTTTTGTTTTTTGTTTTTGAGACAGA 621

Qy    628 GTCTCAACTCTGTCACCCAGGCTGGAGTGCAGTGGCGCGAAATCTCGGCTCACTGCAACCTC 687
          | ||| | |||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    620 GTCTC-ACTTTGATGCCCAGGCTGGAGTGCAGTGGCTCACTGCAACCTT 562

Qy    688 TGCCTCCCGGGTTCAAGTTATTCTCCTGCCCCAGCCTCCTGAGTAGCTGGGACTACAGGC 747
          |||||||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db    561 CGCCTCCCAGGTTCAAGCAAATTCCCTGCCTCAGCCTCCCAACTAGCTGGGACTACAGGT 502

Qy    748 GC---CCACCACGCTAGCTAATTTTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGT 804
          ||   |||||||||||||||| ||||| |||||| |||||||||| ||||||||||||||
Db    501 GCCTGCCACCACGCCCAGCTAATTTTTTATATTTTTAGTAGAGACA-GGGTTTCACCATGT 443

Qy    805 TCGCCCAGGTTGATCTTGATCTCTGGGACCT--TGTGATCTGGCCTCGCCTCGGCCTCCCAAAG 862
          |||||||||||||||||||||||| ||||||  |||||||  ||| |||||||||||||||||
Db    442 TGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCGGCTAGCCTTGGCCTCCCAAAG 383

Qy    863 TGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTATTTTTTAATTTTTTTGTTTTGA 922
          ||||||||||||||||||| |||||||||||| ||||| ||||||||||||||||||||||
Db    382 TGCTGGGATTATAAGCGTGAGCCACCACGCCCAACCAGTAATTTTTTTTTTTTTTG 323

Qy    923 AATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAATGGCCAAATCTCGGCTCACTGCA 982
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db    322 GATGGAGTCTCACTCTGTTACCCAGGCTGGAGTGCAGTGGCCACGATCTTGGCTCACCGCA 263

Qy    983 ACCTCTGCCTCCCGGGCTCAAGCGAATTCTCCTGTCTCAGCCTCCCAAGCGACTGGGATTA 1042
          ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    262 ACCTCTGCCTCCCAGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTA 203

Qy   1043 CGGGCACCTGCCACCACACCCGGCTAATTTTGTATTTTTCATTAGAGGCGGGGGGTTTCACC 1102
          |||||||||||||||||||||| |||||||| ||||||||| |||||||| ||||||||||
Db    202 CGGGCACGTGCCACCACGCCCGGCTAATTTTTGTATTTTTAGTAGAGACAGGGGTTTCGCC 143

Qy   1103 ATATTTGTCAGGCTGGTCTCAAACTCCTGACCTCAGGTGACCCAGGCTGCTTCCGCC 1162
          ||| ||| |||||||| || ||||||||||||||||||||| |||||||||||||||
Db    142 ATGTTGTTCAGGCTGGTCTCGAACTCCTCACCTCAGGTGATCTGCCTACCTCAGCCTCCC 83
```

```
Qy    867 TTTTATTTTTAATAGAGATGTCTCACTATGTTGCCCGGGACTAGTCTCAAAACTCCTGGGC 926
          ||||| |||||  |   |||||||||||||||||||||||||||||||||||||||||||
Qy    361 TCAAGCAGTCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAG--GCGTGCAGCCGT 418
          |||||  ||| |  |||||||||||||||||||||||||||||||||  ||||||||||||
Db    927 TCAAGCCATCCCACCTT-GGCCTCCCAAAGTGCTGGGATTACAGTTCGAGCCCACTGT 985

Qy    419 GCCTGGCCTTTTATTTTTTATTTTTTTAGACACAGGGTCTCACCGAGGATGA 478
          |||||||  ||||||||| ||||| |||| ||||||||||||||| |||||
Db    986 GCCTGGCCTCTTTTTTTTTTGAGATAGA-----ATCTCCCTGTCACCAGGTTGG 1038

Qy    479 AGTGCAGTGGTGTGATCACAGCTCACT------GCAGCCTTCAAC 517
          ||||||||||||||  |||||||||||      ||||||||||||
Db   1039 ACTGCAGTGGTGCGATCTCAGCTCACTGCAACCTCCGCCTCCGTC 1098

Qy    518 TCCTGAGATCAAGC-ATCCCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACCAAAGACATGC 576
          |||||||||||||| || ||||||||||||||||||||||||||||||||||  | |||||
Db   1099 TCCTGGGTTCAAGCAAATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCAT 1158

Qy    577 ACCACTACACCTGCTAATTTTTTATTTTTTTTAATTTTTTTGAGACAGAGTCTCAACT 636
          |||||| || ||||||||| ||| ||||||||| |||||||||||||||||||||| ||
Db   1159 GCCACCACGCCTGGCTAATTTTGTGTTTTTTTAGAGACGAGTCTC---CT 1215

Qy    637 CTGTCACCCAGGCTGGAGTGCAGTGGCGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCG 696
          ||||||||||||||||||||||||||| || |||||||||||||||||||||||||||||||
Db   1216 CTGTCA-CCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAACCTCCGCCTCCCG 1274

Qy    697 GGTTCAAGTTATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTACAGGCGCCCACCAC 756
          ||||||||  || ||||||||||||||||||||||||||||||||||||||||||||||
Db   1275 GGTTCACGCCATTCTCCTG--GCCACCACGCCCAGCTAGGCTGGGACTACAGGCGCCCAC 1333

Qy    757 GCCTAGCTAAT--TTTTTTGTATTAGAGAGAGGG--TTCACCATGTTGTCCCAGGTT 814
          |||| |||||   ||| |||||| |||||| |||  |||||||||||||  |||||||
Db   1334 CATGCCCGGCTAATTTTGTATTTTTAGTAGAGAGGGGTTTCACCATGTTGGCCAGGAT 1393

Qy    815 GATCTGATCTCTGGGACCTTGTGATCTGCGGCTCCCAAAGTGCTGGGATTAC 874
          |||| ||||||||||| || ||||||| |||| |||||||||||||||||||
Db   1394 GGTCTCAAACTCCTGACCTCATGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTAC 1453

Qy    875 AGGGGTGAGCCACCACGCCCGGCCTTAATTCTTTGTTTTGTTTTTGTTTGAAATGGAATCTCA 934
          ||| ||||||||||||||||||| ||||||||||| ||||||||| |||| ||||||||||||
Db   1454 AGGTGTGAGCCATCACGCCCAGCCTTTATTTTTGTTTTTTTT---TAGATCTACC 1504

Qy    935 CTCTGTTACCCAGGCTGGAGTGGCGCAAATGGCGTCAATGGCGTTCTCAGCCTCCTGCCTCC 994
          ||||||||||||||||||||||  ||||| |||||| | | |||||||||||||| |||||
Db   1505 CTCTTGTTGCCCAGGGGAGTCCAGCCTTGGCTGCTCCAGCCTAACCGCGGCGCCTCC 1564

Qy    995 CGGGCTCAAGCGAATTCTCCTGTCTCTCAGCCTCCCAAGCCAGCT 1035
          ||||||||||||||| ||||||||||||| ||||||||||||||
Db   1565 CGGGTTCAAGCGAATTCTCCTGCCTCCGAGTGTCCCGAGTGTCT 1605
```

RESULT 4

```
Qy 1163 AAAGTGCTGGGATTAYACAGGCGGTGAGCCACCTCACCCAGCCGGCTAAT 1209
        |||||||||||||||| ||||| ||||| ||||||||    |||   |||
Db   82 AAAGTGCTGGGATTATAYAGGCATGAGCCACCACATCTGGCCCAGTAAT    36
```

```
LOCUS       AV762220             1345 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AV762220 MDS Homo sapiens cDNA clone MDSCBD07 5', mRNA sequence.
ACCESSION   AV762220
VERSION     AV762220.1  GI:10920068
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1345)
  AUTHORS   Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
            Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
            ,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
            Yang,Y., Gao,G., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
  TITLE     Unpublished (2000)
  JOURNAL   
  COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1..1345
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="MDSCBD07"
                     /clone_lib="MDS"
                     /tissue_type="Bone marrow"
                     /cell_type="CD34+ hematopoietic stem/progenitor cell"
                     /lab_host="BM25.8"
                     /note="vector:pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      388 a    316 c    354 g    277 t    10 others
ORIGIN
```

Query Match 26.6%; Score 383.6; DB 10; Length 1345;
Best Local Similarity 82.8%; Pred. No. 3.5e-61;
Matches 502; Conservative 0; Mismatches 86; Indels 18; Gaps 5;

```
Qy  600 ATTTTTAATTTTTAATTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAG 659
        ||||||   |   |  |||||||||| |||||| ||||| |||| |||||| |||||||||
Db  606 ATTTTTNTNTNTTTTTTTTTGAGACGCAGTCTC-ACTCYGTCGCCCAGGCTGAAGTGCAG 548

Qy  660 TGGCGCAATCTTGGCTCACTGCAACCCTCTGCCTCCCGGGTTCAAGTYATTCTCCTGCCCC 719
        ||||| | ||| ||||||||||||||||| | | |||  ||||||| ||||||||||||| |
Db  547 TGGCGTGATCTCAGCTCACTGCAACCCTCCGTGGTTCAAGCAATTCTCCTGCCTC       488

Qy  720 AGCCTCCTGAGTAGCTGGGACTACAGGCGC--CCACCACGCGTYAGCCTAATTTTTTTGTA 776
        |||||||||||||||||||||||||||||||  |||||||||| | ||||||||||| |||
Db  487 AGCCTCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCGGCTAATTTTTGTATT   428

Qy  777 TTTTTAGTAGAGATGGGGTTTCACCATGTTGCCAGGTTGATCTCGGGACCTTGT        836
        ||||||||||||| | |||||||| |||||| ||||| |||||| |  |||  |||
Db  427 TTTAGTAGAGACA-GGGTTTCATCATGTTGGCCAGGCTGGTCTTGAAACTCCTGACCTCGT 369

Qy  837 GATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGAATTACAGGCGTGAGCCACCACGCCCG  896
        ||||  ||  ||||||||||||||||||||||||| |||||||||| ||||||||  ||||
Db  368 GATCCACCACGCGACTCCTCAAGTGCTGGGATTACAGGCGTCAGCCACCATGCCCG      309

Qy  897 C-----------TTATTTTTAATTTTTGTTTGTTTGAAATGGGAANTCTCACTCTGTTACCC 945
        |            |||||| |  ||||| ||||||| ||||| ||||||||||||| ||||
Db  308 CCTCATTTGCCATTCTTTTTTTTTTTTTCTTTTTTTTGAGATGGAGTCTCACTCTATTGCCC 249
```

RESULT 5

```
Qy 1186 AGCCAC 1191
        | ||||
Db   10 AACCAC    5
```

```
LOCUS       BQ722917              910 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8100507 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6190443 5', mRNA sequence.
ACCESSION   BQ722917
VERSION     BQ722917.1  GI:2186 1803
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 910)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
            cDNA Library Preparation: Dr. James R. Lupski
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13589 row: e column: 04
            High quality sequence stop: 653.
FEATURES             Location/Qualifiers
     source          1..910
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6190443"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCACCGGTCCG-3' and
                     5'-GACTACGTTCTAGATCCGGAGCGCGGCCGCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.9 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine); available through Life
                     Technologies."
BASE COUNT      277 a    224 c    268 g    140 t    1 others
ORIGIN
```

```
Qy  946 AGGGCTGGGAGTGCAAATGGCCAAAATCTCGGGCTCACTCACTGCAAGGCGCTGCCTCAGC 1005
        ||| ||||||||| |  ||| |||||||||||||||||  |||||| | || ||| || |||
Db  248 AGGCTGGAGTGCTGTGGCGCAATCTCGGGCTCACTGCAACCCTGCCCTCCCGGGTTCAAGT 189

Qy 1006 GATTCTCCTGCGTCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCCCGCCACCACCCCG 1065
        ||||||||||| |||||||||||| ||||||||||||||||||||||| |||||||| |||
Db  188 GATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCCG  129

Qy 1066 CTAATTTTTGTATTTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGGTCTCAAA 1125
        |||||||||||||||||| ||||||| |||||||||||| ||| | |||||||||||| ||
Db  128 CTAATTTTTGTATTTTTACTAGAGACGGGGTTTCACCATGTTGAGGCTGGTCTCGAAA    69

Qy 1126 CTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG 1185
        |||||||||||  |||||||| |||||||||||||||||||||||||||||||||||||||
Db   68 CTCCTGACCTC--GTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTG   11
```

Query Match 26.3%; Score 379; DB 14; Length 910;
Best Local Similarity 80.8%; Pred. No. 2.8e-60;

Matches   514; Conservative   0; Mismatches   91; Indels   31; Gaps   5;

Qy   595 TTTTTATTTTATTTTTTAATTTTTTGAGACAGAGTCTCAAACTGTCTGTCACCAGGCTGGAG 654
         ||||| ||||||||||| |||||| ||| | |||||||| | | ||||||| | |||||||
Db   643 TTTTT--TTTATTTTTTTTAATTTTTTTTTTTTTTTGAGGCGGAGTCT-TGCTCTGTCGCCCAGGCTGGAG 585

Qy   655 TGCAGTGGGGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCT 714
         ||||||| ||||||||||||||||||||||||||||||||||| |||||||| |||||||||
Db   584 TGCAGTAGCGCAAATCTTGGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGCGATTCTCCT 525

Qy   715 GCCCCAGCCTTCCTGAGTAGCTGGGACTACAGGCGCC--CCACCACGCCCAGCTAATTTTTT 771
         ||| ||||||||||||||||||||||||||||||| |||||||| | ||||||||||
Db   524 GCCTCAGCCTTCCTGAGTAGCTGGGACTACAGGCACACATGCCACCACGTCCGGCTAATTTTT 465

Qy   772 TTTGTATTTTTTAGTAGAGATGGGGTTTCACCATGTTCGCCCAGGCTGGTCTCGATCTCCTGAC 831
         |||| | |||| ||| ||| ||||||||||||||| |||||||||||||||||||||||||| |
Db   464 TGTATTTTTAATGGAGACAGGGTTTCACCATGTTGCCCAGGCTGGTCTCGATCTCCTGAT 405

Qy   832 CTTGTGATCTCGCCCGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG 891
         |||||||||| |||| ||||| |||||||||||||||||||||||||||||||||
Db   404 CTTGTGATCTGCCCTCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG 345

Qy   892 CCCGGC--TTATTTTTAATTTTGTTTTGTTTTGAAATGG 927
Db   344 CCCGGCCCCCCTAGTATTTCTTTGTTTTGTATTTTCAGTAGCTCCTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG 285

Qy   928 AATCTCACTCTGTTACCAGGCTGGAGTGCAGTGGCCAATTCTCGGCTCACTGCAACCTC 987
         |||||||||| ||| |||||||||||||| ||||||| |||||||||||||||||||||
Db   284 AGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTC 225

Qy   988 TGCCTCCCGGGCTCAAGCGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACGGGGC 1047
         |||||||| |||||||| ||||||||||||||||||||| ||||||||||||||| || |
Db   224 TGCCTTCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGC 165

Qy 1048 ACCTGCCCACCACGCCCGGCTAA-TTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATAT 1106
         |||||||||| ||| |||||||| || | |||||||||||||||| ||||||||||| |
Db   164 GCCCGCCACTACCGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTTT 105

Qy 1107 TTGTCAGGCTGGTCTCAAAACTCCTGACCTCAGGTGACCCGCCCGCCTTCAAAG 1166
         | |||||||||| |||| |||||||||||||||||| ||| ||||||| |||||
Db   104 TAGCCAGGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCCCAAAG 47

Qy 1167 TGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCC 1202
         |||||||||||||||||||||||||||||||||| |
Db   46 TGCTGGGATTACAGGCGTGAGCCACCATGCCCGGCC 11

RESULT 6
LOCUS   AV700498
DEFINITION AV700498 GKC Homo sapiens cDNA clone GKCAAC05 3', mRNA sequence.
ACCESSION   AV700498
VERSION   AV700498.1 GI:10302469
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 767)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z. and Han,Z.
  TITLE   Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
  COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            20123, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.

LOCUS   AV700498   767 bp   mRNA   linear   EST 16-JAN-2002

FEATURES   Location/Qualifiers
     source   1..767
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GKCAAC05"
                     /clone_lib="GKC"
                     /tissue_type="hepatocellular carcinoma"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT   148 a   218 c   184 g   216 t   1 others
ORIGIN

Query Match   26.2%; Score 378.4; DB 10; Length 767;
Best Local Similarity 83.1%; Pred. No. 3.8e-60;
Matches 481; Conservative   0; Mismatches   87; Indels   11; Gaps   4;

Qy   635 CTCTGTCACCCCAGGCGTGGAGTGCAGTGGCGGGCCAATCTTGGCTCACTGCAACCTCTGCCTCC 694
Db   13 CTCTGTCACCCCAGGCTGGAGTGCAGTGGCGCGATCTCGGCTCACTGCAAGCTCTTCCTCC 72

Qy   695 CGGGTTCAAGTTATTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACGGGGCAGCTCCA 751
Db   73 CGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCCCGCCA 132

Qy   752 ACCACGCCTAGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTCGCCA 810
Db   133 ACCACGCCCGGCTAATTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCA 192

Qy   811 GGTTGAGTCTTGGATCTCCTGACCTTGTGATCGTCCGCCTCGGCCTCCCAAAGTGCTGGGA 870
Db   193 GGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGA 252

Qy   871 TTACAGGCGTGAGCCACCACGCCCGGCCTATTTTTGTATTTTTCATTAGAGGCGGGGTTTCACCATATTTTG 930
Db   253 TTACAGGCGTGAGCCACCGGCCTCT-----TATTTTTTTTTTTGTTTTGTTTTGAAATGCTGG 306

Qy   931 CTCACTCTGTTACCAGGCTGGAGTGCAAATGGCGAAATCTCGGCTCACTGCAACCTCTGC 990
Db   307 CTCACACTGTCGCCCAGGCTGGAGTGCAGTCGGCGATCTCGGCTCACTGCGAACCTCCGC 366

Qy   991 CTCCCGGGCTCAAGCGATTCTCCTGTCTCAGCCTCCCAAAGCACGTGGGATTACAGGCTGC 1050
Db   367 CTCCTGGGTTCAAGCGAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGGCGCC 426

Qy 1051 TGCCACCACACCCCGCTA-ATTTTTGTAGTTTTCATTAGAGGCGGGGTTTCACCATATTTG 1109
Db   427 CACCACACGCCCTTAGTTTTGTTTTTGTATTTTTAGTAAAGATGGGGTTTCACCATGTTGG 486

Qy 1110 TCAGGCTGGTCTCAAAACTCCTGACCTCAGGTGACCCGCCCGCCTTCAAAGTGC 1169
Db   487 CCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCCGNCCACCTTAGCCCTCCCAAAGTGC 546

Qy 1170 TGGGATTACAGGCGTGAGCCACCTCACCCAGGCCGGCTAA 1208
Db   547 TTGGATTACAGGCGTGAGCCACCATACCTGGCCAGCAAA 585

RESULT 7
LOCUS   AQ393450
DEFINITION CITBI-E1-2546H6.TF CITBI-E1 Homo sapiens genomic clone 2546H6, DNA
            sequence.
ACCESSION   AQ393450
VERSION   AQ393450.1 GI:4364473
KEYWORDS   GSS.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

LOCUS   AQ393450   658 bp   DNA   linear   GSS 06-MAR-1999

```
REFERENCE   1 (bases 1 to 658)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE       Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
            Map Building
JOURNAL     Unpublished (1997)
COMMENT     Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13-21
            CalTech Human BAC Library D"
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..658
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2546H6"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
                     /note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
BASE COUNT    106 a    183 c    163 g    206 t
ORIGIN

Query Match   25.9%;  Score 373.8;  DB 17;  Length 658;
Best Local Similarity   80.9%;  Pred. No. 2.8e-59;
Matches 512;  Conservative  0;  Mismatches 107;  Indels  14;  Gaps  6;

Qy  596 TTTTTATTTTTTATTTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGT 655
        ||||  |||||||||||| |||||| |||||||||||||||||||  |||||||||||||||||
Db    8 TTTTT.TTTTTTTTTTATTTTTTTTTTGAGACGGAGTCTC.ACTCTGTCACCCAGGCTGGAGT 66

Qy  656 GCAGTGGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAGTTATTCTCCTG 715
        |||||||  |  |||| ||||||||||||||||| | |||||||| ||||||||||||||
Db   67 GCAGTGGCGATGATTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGGATTCTCCTG 126

Qy  716 CCCCAGGCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCGGCTAGTTAAT-TTTTTG 774
        |  |||||||| |  |||||||||| || |||| |||| ||||| |||||  |||| || |||||||
Db  127 CCTCAGCCTCCCAAAGTAGCTGGGACTAGAGGTGCGCCCACCACGCCCGGATAATTTTG 186

Qy  775 TATTTTTTAGTAGAGATGGGG-TTCACCATGTTGGCCAGGTTGATCTTGATCTCTGGACCT 833
        |||||| |||||||||||||| ||||||||||||| |||||||| | |  ||||| | |
Db  187 TATTTTTAGTAGAGATGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGAATCTCTGACCT 246

Qy  834 TGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCAGCGCC 893
        ||| || |||||||||| |||||||| |||||| |||||||||||| ||||||||| | |||
Db  247 CGTAATTCTCCTGCCTTGGCCTCCCAGAGTGCTGGGATTACAGGCGTGAGCCACCACTGCCC 306

Qy  894 CGGC----------TTATTTTAATTTTTGTTTTGTTTTGTTTTGAAATCTCACTC-TGTTAC 943
        ||||          ||||||  ||||||||||||||||||  |||||||||  ||| |||| |||
Db  307 CAGCCCCGGCCCTCTTTTTTTTTTTTTGTTTTGTTTTGTTTTTCGAGGGGGGAGTTTCGGTTCTTGTTGC 366

Qy  944 CCAGGCTGGAGTGGCAAATGGCCAAAATCTCGGCTCACTGCAAGCTCCGCCTCCCGGGTTCAA 1003
        |||||||||||||   ||| |||  ||||||||||||||||| | | |  |  ||| |||||||
Db  367 CCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAA 426

Qy 1004 GCGATTCTCCTGTCTCAGCCTCCCAAGCACCTGGGATTACAGGCGCCCACCACCACACC 1063
        | |||||||||| ||||||||||||| || || |||||||| |||||| |||||||| |||
Db  427 GTGATTCTCCTGCCTCAGCGTCACGAGTAGCTGGGATTACAGGCGTCCACCACCACGCC 486

Qy 1064 CGCTAA-TTTTTGTATTTGTATTTTTCATTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCTC 1122
        ||| || ||||||||||| ||||||| || || || ||||||||||| || || ||||||||||||
Db  487 AGTCAATTTTTTGTATTTTTTAGTAGAGTAGGGTTTCACCATTTTGGCCAGGCTGGTCTC 546

Qy 1123 AAACTCTCGACCTCAGGTGACCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGC 1182
        |||||| | || ||||||||  |||| |||| ||||||||||||||||| ||||||||||
Db  547 GAACTCCTGACCTCAGGTGATCAGGTGATCTAGCTGGCCGCCTCCCAAAGTGCTAGGATTATAGGT 606

Qy 1183 GTGAGCCACCTCACCCAGCCGGCGGCTAATTAGAT 1215
        |||||||||||  |||| |||| |  ||| ||| |
Db  607 GTGAGCCACCACACCCGGCCCGCCCCCATTTTGTAT 639
```

```
                                                               Page 6

RESULT 8
AQ746594

LOCUS       AQ746594   911 bp   DNA   linear   GSS 19-JUL-1999
DEFINITION  HS_2278_A2_E05_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2278 Col=10 Row=I, DNA sequence.
ACCESSION   AQ746594
VERSION     AQ746594.1  GI:5533752
KEYWORDS    GSS.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 911)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE     99380589
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2278  row: I  column: 10
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 911.
FEATURES             Location/Qualifiers
     source          1..911
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=2278 Col=10 Row=I"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT    193 a    237 c    245 g    235 t    1 others
ORIGIN

Query Match   25.9%;  Score 373;  DB 17;  Length 911;
Best Local Similarity   79.1%;  Pred. No. 3.6e-59;
Matches 493;  Conservative  0;  Mismatches 125;  Indels  5;  Gaps  4;

Qy  599 TATTTTTATTTTTTAATTTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCA 658
        ||||| ||| ||||||||||||||||| ||||||||| |||| | |||||||||||||||||
Db   42 TTTTTGTTTGTTTTGTTTTGTTTTGAGATGGAGTCTC.ACTCTTTCACCCAGGCTGGAGTGCA 100

Qy  659 GTGGCGCGAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCC 718
        |||||||  | ||| ||||||||||| | |  || | |||| ||||| |||| |||||||||
Db  101 GTGGCGCAGTCTCGGCTCACTGCAAGCCCCACCTTCTCGGTTCATTACCCATTCTCCTGCCT 160

Qy  719 CAGCCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCACGCCCACCTAGCTAATTTTTTGTATT 778
        || ||||| |||||||||||||||||||||||| |||||||||||  ||| |||||||||||||
Db  161 CAACCTCCCTGAGTAGCTGGGACTACAGGCGCCCGCCACCACGCCCGGCTAATTTTTTGTATT 220

Qy  779 TTTTAGTAGAGA-TGGGGTTTCACCATGTTGCCCAGGTTGATCTTGATCTCTGGACCTTGTG 837
        ||| ||||||||| |||||||||||| |||||| || |||  || |||| | | ||| |||||
Db  221 TTTAGTAGAGACGGGGTTTCACCTGTTGTTAGCCCAGGATGGTCTGGATCTCCTGACCCTGTG 280
```

```
Qy    838  ATCTGCCTGCCTCCTCCTCCCCTCCCCAAAGTGCTGTGGGATTACAGGCGTGAGCCACCACGCCCGGC  897
           ||||| | |||||||||| ||||| ||||| ||||||||||||||||||||||||||||| |||||||
Db    281  ATCCACCACGCCCTCGGCCTTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGTTCCCGGC  340

Qy    898  --TTATTTTTAATTTTTGTTTTGTTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGT  955
           |||||||| |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    341  TGTTTGTTTTTATTTTGTTTTGTTTGAGACAGGGTCTCCCTCTGTTGCCCAGGCTGGAGT  400

Qy    956  GCAATGGCCAAAATCTCGGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTG  1015
           |||||||||||| ||||| |||||||||||||||||||||||||||||||||| ||
Db    401  GCAGTGGCA--AGCTCTCGGCTCACTGCAACCACTGCCTCCCAGGGTTCAAGCGATTCTCTG  459

Qy    1016 TCTCAGCCTCCCAAGCAGCAGTGGGGATTACAGGCGGGCGCCACCACGCCCGGCTAATTTTTG  1075
           || ||||||||| |||||||||||| |||||||||||||||||||||||||||||||| |||
Db    460  CCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGTGCGCCACCACGCCCGGCTAATTTTTG  519

Qy    1076 TATTTTTCATTTAGAGACGGGGTTTCACCATGTTGTCAGGCTGGTCTCTCAAACTCCTGACCT  1135
           |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db    520  TATTTTTAGTAGAGATGGGGTTTCACCATGTTGTCCAGGCTGGTCTCTCAAACTCCTGACCTT  579

Qy    1136 CAGGTGACCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCTCA  1195
           | |||||||| ||||||||||||||||||||||||||||||||||||| ||||||||
Db    580  CAAGTGATCCACCTGCCTTGGCCTCCCAAAGTGTTGAGATTACAGGAGTGAGCCATTGCG  639

Qy    1196 CCCAGCGCGGCTAATTTAGATAAA  1218
           ||||||||||||||||||||||||
Db    640  CCCGGCAGCAAAAGTTTAAACA  662
```

```
RESULT 9
LOCUS       BQ438614/c
DEFINITION  BQ438614        637 bp    mRNA     linear
            AGENCOURT_7906616 NIH_MGC_82 Homo saplens cDNA clone IMAGE:6104226
            5', mRNA sequence.
ACCESSION   BQ438614
VERSION     BQ438614.1  GI:21177690
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2341 row: 1 column: 19
            High quality sequence stop: 526.
FEATURES             Location/Qualifiers
     source          1..637
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="IMAGE:6104226"
                     /clone="IMAGE:6104226"
                     /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                     Sfil (ggcgcctcggcc); Site_2: Sfil (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence: 5'-CACGGCCATTATGGCC-3', and 3' adaptor sequence:
                     5'-APTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C, G, or T). Average insert size
                     1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained
                     inserts by PCR. This library was enriched for full-length
                     clones and was constructed by Clontech Laboratories (Palo
                     Alto, CA)."
```

```
BASE COUNT      178 a     172 c     182 g     102 t      3 others
ORIGIN

Query Match     25.7%;  Score 370;  DB 14;  Length 637;
Best Local Similarity   79.4%;  Pred. No. 1.4e-58;
Matches  494;  Conservative  0;  Mismatches  107;  Indels  21;  Gaps  4;
```

```
Qy    596  TTTTATTTTATTTTTAATTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGT  655
           |||| | |||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    621  TTTTTCTTTTTCTTTTTTTTTTTTTTNGGAGATGGAGTCTC-GCTCTGTCGCCCAGGCTGGAGT  563

Qy    656  GCAGTGGCGCCAAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGATTCTCCTG  715
           |||||||||||||||||||| |||||||||||||||||||| ||||||||||||||||||
Db    562  GCAGTGGCGACGATCTTGGTTCACTGTAAGCTCCGCCTCCTCNCAGGGTTCAAGTTATTCTCCTG  503

Qy    716  CCCCAGCCTCCTGATTAGCTGGGACTACAGGCGCG----CCACCACGCCCTAGCTAATTTTTT  772
           ||||||||||||| |||||||| |||||||||  ||||||||||||||| |||||||||
Db    502  CCTCAGCCTCCCAAGTAGCTGGGACTACAGGCGCCCACCATGCCCAGCTAATTTTT  443

Qy    773  TGTATTTTTAGTAGAGATGGGG-TTCACCATGTTCGCGCCAGGTTGAATCTCTGGAC  831
           ||||| |||||||||||||||| ||||||||||| | ||| ||| || | ||||||
Db    442  TGTATTTTTAGTAGAGACGGGGTTTTTACCGTGTTAGCCAGGATGGTCTCGACCTCCTGAC  383

Qy    832  CTTGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACG  891
           || |||||| || |||||||||||||||||||||||||||||||||||||||||||||||
Db    382  CTCGTGATCTGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCG  323

Qy    892  CCGG-------------------CTTATTTTTAATTTTTGTTTTGTTTTGAAATGGAATCTCAC  935
           ||||                   |||||||||||||||||||||||| |||||||||||||||
Db    322  CCCGGCCGGGAAGCATGTTTCTGAAAACATAACCTAAGAAAAGTATAAAATGCTGTGCTTGCT  263

Qy    936  TCTGTTACCCAGGCTGGAGTGCAATGGCCAAAATCTCGGCTCACTCTGCCTCCC  995
           ||||||||||||||||||| ||||||| |||||||||| ||||||| |||||||
Db    262  CTTGTCACCCAGGCTGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCCGCCTCCC  203

Qy    996  GGGCTCAAGCGAITTCTCCTGTCTCAGCCTCCCAAGCAGCAGTGGGGATTACAGGCACGCCA  1055
           |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    202  GGGTTCAAGTGAITTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCGTGCCA  143

Qy    1056 CCACACCCCGGCTAATTTTTGTATTTTTCATTAGAGACGGGGGTTTCACCATATTGTTCAGGC  1115
           |||| | |||||||||| ||||||| ||||||||||||||||| |||||||||| |||||
Db    142  CCATGCCCGGCTAATTTTTGTATTTTTGTAGAGATGGGGTTTCACCGTGTTGGCCAGGC  83

Qy    1116 TGGTCTCAAACTCCTGACCTCAGGTGACCACCCTGCCTCGGCCTCCCAAAGTGCTGGGAT  1175
           ||||| ||||| |||||||||||| |||||||| | || ||||||||||||||||||||
Db    82   TGGTCTGGAACTCCTGATCTCGGGTGATCCGTCTGCCTCGGCCTCCCAAAGTGCTGGGAT  23

Qy    1176 TACAGGCGTGAGCCACCTCACC  1197
           |||||||| | |||||||||||
Db    22   TACAGGCGTGAGCCACCGCC  1
```

```
RESULT 10
LOCUS       BM468547/c
DEFINITION  BM468547        970 bp    mRNA     linear
            AGENCOURT_6475537 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578163
            5', mRNA sequence.
ACCESSION   BM468547
VERSION     BM468547.1  GI:18517589
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

EST 24-MAY-2002

EST 05-FEB-2002

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12332 row: m column: 12
High quality sequence stop: 671.

```
FEATURES
     source
                     1..970
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5578163"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2 kb. Library constructed by Life
                     Technologies."

BASE COUNT     291 a    235 c    240 g    195 t    9 others
ORIGIN

Query Match   25.6%;    Score 368.8;    DB 13;    Length 970;
Best Local Similarity  78.3%;    Pred. No. 2.1e-58;
Matches 519;   Conservative 0;   Mismatches 133;   Indels 11;   Gaps 6;

Qy   566 CAAAGACAATGCACCACTACACCTGGCTAATTTTTTTATTTTTATTTTTAATTTTTTGAGACA 625
Db   904 CAAATCGAGTCTCCCCAAGGCATTGCGGGATCAGAATTTTTTTTTTTATTTTTTGTATA 845

Qy   626 GAGTCTCAACTCTGTTCACCCAGGCTGGAGTGCAGTGGCGCAAATCTTGGCTCACTGCAACC 685
Db   844 GAGTCTCCGC-CTGTCTCCCAGGCTGGGAGTGCAGTGGCGGGATCTCAGCTCACTGCAAGC 786

Qy   686 TCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAG 745
Db   785 TCCGCCTTCTGGATTCCCGCCACCATTCTCCTGCCTCAGCCTCNTGAGGTAGCTGGGACTACAG 726

Qy   746 GCGCCCACCA---CGCCTAGCTAATTTTTTTTTTTGTATTTTTTTAGTAGAGACA---TGGGGTTTCAC 799
Db   725 GCGCCCGCCACTGCACCCGCCTAATTTTTTGTATTTTTTTAGTAGAGACGGGGTTTCAC 666

Qy   800 CATGTTCGCCAGGTTGATCTGATCTTGTGATCGTCGTGCCTCGCCTTGGCCTCCCA 859
Db   665 CGTGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTGATCCGCCCAGCCTCGGCCTCCCA 606

Qy   860 AAGTGGCTGGGATTACAGGCGTGAGCCACCACGCCCCGGCCTTATTTTTTGTTTTTTGTT 919
Db   605 AAGTGCTGGGAATTACAGGCATGAGCCACCGCCGCCGGCCTTTTTTTT-TTTTTTCTTTTTTA 547

Qy   920 TGAAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGCAAATGGCGCAAATCTCGGCTCACT 979
Db   546 AGACATAGTTTCCGTCTT-TGTCGCCCAGGCTGGAGTGCATGGCGCGATCTCGGCTCACC 488

Qy   980 GCAACCTCTGCCTCCCGGGCTCAAGCGAATTCTCGTCTCCCAAGCCAGCTGGGA 1039
Db   487 GCAACCTCCGCCTCCTGGGGAATTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGA 428

Qy  1040 TTACGGGCGACCTGCCACCACACCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGGTTTC 1099
Db   427 ATACAGGCATGGTGCCACCATGCCCAGCTAATTTTTTGTATTTTTAGTAGACGGGGTTTC 368

Qy  1100 ACCATATTGGTCAGGCTGGTCTCAAACTCCTGACCTCGTGATCCTGCCTCAGCCT 1159
Db   367 ACCATCTTGGCCAGGCTGGTCTCGAACTCCTGACCTC--GTGATCCACCCGCCTCGGCCT 310

Qy  1160 TCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTAATTTAGATAAAA 1219
Db   309 CCCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCACGGGGATCAGAGTTTTTA 250

Qy  1220 AAA 1222
Db   249 AGA 247
```

RESULT 11
AQ839814

```
LOCUS       AQ839814             5556 bp    DNA     linear   GSS 31-MAR-2000
DEFINITION  260L13-C50 CITB Homo sapiens genomic clone 260L13, DNA sequence.
ACCESSION   AQ839814
VERSION     AQ839814.1  GI:6652446
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5556)
  AUTHORS   Carpten,J.D., Makalowska,I., Robbins,C.M., Scott,N., Sood,R.,
            Connors,T.D., Bonner,T.I., Smith,J.R., Faruque,M.U., Stephan,D.A.,
            Pinkett,H., Morgenbesser,S.D., Su,K., Graham,C., Gregory,S.G.,
            Williams,H., McDonald,L., Baxevanis,A.D., Klingler,K.W. and Landes
            ,G.M.
  TITLE     A 6-Mb high-resolution physical and transcription map encompassing
            the hereditary prostate cancer 1 (HPC1) region
  JOURNAL   Genomics 64 (1), 1-14 (2000)
  MEDLINE   20175426
COMMENT     Contact: Carpten JD
            Cancer Genetics Branch
            National Human Genome Research Institute/National Institutes of
            Health
            Bldg. 36, Room 3D04, 36 Convent Drive, Bethesda, MD
            Tel: 301 435 5626
            Fax: 301 435 5465
            Email: jdc@nhgri.nih.gov
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..5556
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="260L13"
                     /clone_lib="CITB"
                     /note="Kim, et al., Genomics, 34, 213"

BASE COUNT    1500 a    1016 c    1002 g    2035 t    3 others
ORIGIN

Query Match   25.5%;    Score 367.6;    DB 17;    Length 5556;
Best Local Similarity  80.2%;    Pred. No. 1.9e-58;
Matches 483;   Conservative 0;   Mismatches 109;   Indels 10;   Gaps 4;

Qy   604 TTATTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGTGGC 663
Db  3145 TTAACTTTTTTTTTTTTTTGAGATGGAGTCTCGCTCTGTCGCCCAGGCTGGAGTGCAGTGGC 3204

Qy   664 GCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCTGCCTCAGCC 723
Db  3205 GCGATCTCGGCGCTCACTGCAATCCTTCGCCTCCCGGGTTCAAGGCGATTCTCCTGCCTCAGCC 3264

Qy   724 TCCTGAGTAGCTGGGACTACAGGC---CCACCACGCCCAGCTAATTTTTTTTTTGTATTT 780
Db  3265 TCCCGAGTAGCTGGGACTACAGGCGCCCACCACGCCCAGCTAATTTTTTGTATTTTTA 3324

Qy   781 TAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGATCTCCTGGACCTTGTGATC 840
Db  3325 GTAGGGACA-GGGTTCACCATGTTGGCCAGGGATGGTCTCGATCTCCTGGACCTTGTGATC 3383

Qy   841 TGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCCTTA 900
Db  3384 CGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTATAGGCGTGCCCAGCCACCACGTGCCCGGCCTC 3443

Qy   901 TTTTTTAATTTTTTGTTTTGTTGTAGAAATGGAAATCTCACTCTGTTACCCAGGCTGGCAAT 960
Db  3444 TTTTT--TTTTTTTTTTTGTTTGAGATGGGGAGTCTCGCTGTCTCCCAGGCTGGAGTACAGT 3501

Qy   961 GGCCAAATCTCGGCTCACTGCAAACCTCTGCCTCCCGGGCTCAAGCGATTCTCCTGTCTCA 1020
Db  3502 GGCACAATCTTGGCTCGACTGCAACCTTGCCTCCCGGAGTTCAAGCGATTCTCCTGCCTCA 3561
```

```
Qy  1021 GCCTCCCAAGCAGCTGCGGATTACGGCACGCTGCCACCACCCCGCTAATTTTGTATTT 1080
         ||||||||||| ||||||||||| |    ||  ||||| || |||||||||| ||||||
Db  3562 GCCTCCCAAGTAGCTGCGGATTA---CACTCACCACACCTGGCTAATTTTTGTATTT 3617

Qy  1081 TCATTAGAGGCGGGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAACTCCTCAGGT 1140
         |  | |||  |||||| |||||||||| ||  ||||||||||||| |||||| | |||
Db  3618 TTAGTAGAAAATGGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGT 3677

Qy  1141 GACCCACCTGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCACCCAG 1200
         || ||||||||||||||||||||| ||| ||||||||||||||||||||||| ||||| ||
Db  3678 GATCCACCTGCCTCAGCCTCCCAAAAGTGTTGGGATTACAGGCGTGAGCCACCACACCCCG 3737

Qy  1201 CC 1202
         ||
Db  3738 CC 3739
```

RESULT 12
BC024593

LOCUS   BC024593   4087 bp   mRNA   linear   HTC 07-AUG-2002
DEFINITION Homo sapiens, clone IMAGE:3914314, mRNA.
ACCESSION   BC024593
VERSION   BC024593.1 GI:22137609
KEYWORDS   HTC.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 4087)
  AUTHORS   Strausberg,R.
  TITLE   Direct Submission
  JOURNAL   Submitted (01-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK   NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www-shgc.stanford.edu
            Contact: (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 22 Row: m Column: 14
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis
            This clone has the following problem: Incomplete processing.
FEATURES   Location/Qualifiers
     source   1..4087
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3914314"
                     /tissue_type="Uterus, leiomyosarcoma"
                     /clone_lib="NIH_MGC_71"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
BASE COUNT   995 a   861 c   966 g   1265 t
ORIGIN

Query Match   25.5%; Score 367.4; DB 11; Length 4087;
Best Local Similarity   64.2%; Pred. No. 2.3e-58;
Matches 876; Conservative 0; Mismatches 326; Indels 163; Gaps 15;

```
Qy     1 TTTTTTTTTTTTTTGAGATGGAGTTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAACCCGGCAA 60
         ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |
Db  1309 TTTTTTTTTTTTTGAGATGGAGTCTT--GCCCTGTCACCCAGGCTGGAGTGCGGCCA 1366

Qy    61 TCTCAGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCC 120
         |||| |||||||||||||||| || |||||||||||||||||||||||||||||||||
Db  1367 TCTTGGCTCACCGCAAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCCTCAGCC 1426

Qy   121 CAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTATTTTTTAG 180
         |||||||||||||||||||| ||||||||||| |||||||||||||||||||||||||
Db  1427 GAGTAGCTGGGATTACAGGCGTGCCACCACGCCCGGCTAATTTTTGTATTTTTTAG 1486

Qy   181 TAGAGATGGGGTTTCTCATTGTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGGTGATCC 240
          ||||  |||||||||  |||  |||||||||||||||||||||||||| |||||||||||
Db  1487 --AGATAGGGGTTTCACCACTTTGGTCAGGCTGGTCTCGAACTCCTGACCTCAGATGATC 1542

Qy   241 CCTCCGTCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGCCTT 297
         |||| |||| |||||||||||||||||||||||| |||   ||||||| || |||||| ||
Db  1543 CGCCTCCTCGGCCTCCCAAAGTGCCTGGGATTACAGGTGAGAAACCACCGC--GCCACCAGGCCTT 1602

Qy   298 GCCTGGCTCAATTT-------------TTGTGGTAGAAACAGGGTTTCACTGATGTGCCCAAG 346
         ||||| |||| || |             || |||||||||||||||||||| ||||| |||||
Db  1603 TCCTTTTTTTTTTTGAGACGGAGTCTTGCTCTGTCATCCAGGCTTGAGTGTGTGGTGCAG 1662

Qy   347 ------------CTGGCTCCTGAGCTCGTCAAGCGATCCAAGCTCAGCCTCCC 388
                     |||||||||||| |||| |||||||||| ||||||||||||||
Db  1663 TCATGGCTAACGCAGCCTCCGCTGGATCAAGCGATCCCTCGCCTCAGCCTCCC 1722

Qy   389 AAAGTGCTGGGATTACAGGCGTGCAGCCGCCTGCGCCTGGCCTTTTTTTTTTTTTTTTAAG 448
         ||||||| |||||||||||||||| |||| || || ||||||||||||||||||||||  |
Db  1723 GTGTTAGTCTGGGATTACAGGCGTGAGCCACC---ACCATGCCTGGCTAATTTAATTTTTAGT 1779

Qy   449 ACACAGGTGTCCCACTCTTACCCAGGATGAAAGTGCAGTGGTGTGATCAGCTCACTGCA 508
         || || || ||||||| | ||| ||| || || ||||||||||||||||  |||| |||
Db  1780 GACAGGGTCTTGCTATGTTGCTCAGGCTGGTCT-CAAACTCCTGAGCTCAAGCAGACCACCAGCC 1838

Qy   509 GCCTTCAACTCCTGAGATCAAGATCCTCCTGCCTCAGCCTCCCAAAGTAGCTGGGACCAA 568
         |||| |||||||||| || |||  || ||||||||||||||||||||| || |||| |||
Db  1839 CTCCTGCCTTGACCGAAGCAGCCCTCCTCCGGCCTTCCAAAGTGCTGGGATTAC 1898

Qy   569 AGACATGCACCACTACACCTGGC------------------------------------- 591
         || || |||||| ||||||||||
Db  1899 AGGTGTGAGCCACTATACCTAACCCACATGATTTTTTATTTTTTTTGAATATTAGAAATCCTAATAAA 1958

Qy   592 --------------------------TAA 594
                                   |||
Db  1959 ATATACATTAAGAAAAAGTTACAAAAACAAAAAAGAGAAGTTACATTGTGCTTATAC 2018

Qy   595 TTTTTTATTTTTTATTTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGGAG 654
         ||||||| |||| |||||| || |||||||||| |||||||| |||  ||||||||||||||
Db  2019 TTTGAAATTTGTTTTGTTTGTTTTTGAGACAGGGTCT-GGGTCTCGCCACCAGGCTGGAG 2077

Qy   655 TGCAGTGGCGCAATCTTGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTCTCCT 714
         |||||||||||||||||||||||||||||||||| || |||||||||||| ||| ||||
Db  2078 TGTAGTGGTGTGATCATAGCTCAATGCACCAACCTCCGGGCCAAGTGATCGCTC 2137

Qy   715 GCCCCAGCCTCCTCTGAGTAGCTGGGACTACAGGC----------------------GCCCACGCCCAGCTAA 766
         |||||||||||||| |||||||||||| ||||||||                      |||| |||||| ||||
Db  2138 ACCTCAGCCTCCTGGGACTAGCTGGGACTAGGGCTACAGGCATGTGCCACCATGCCCGGCTAATTAATAT 2197

Qy   767 TTTTTTGTATTTTTTAGTAGAGATGCTGGGTTCACCAGGTTGATCTGATCT 825
         |||||| |||||| || | ||||  ||| |||||||  |||| ||| |||| |
Db  2198 ATTTTTTTTATTTTTTGTAGAGATGAGGTTTCATTAGTGTTGCCCAGGCTGGTCTCGGAACT 2257

Qy   826 -CTGGACCTTGTGATCTCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGC 884
          |||||||||||||||| |||| ||||||||||||||||||| ||||| |||||||
Db  2258 CCTGGCCTCAAGTGATCCTCCCGCCTTGGCCTCCCAAAGTGTTGGGATTGGCGTGAGC 2317

Qy   885 CACCACCCCGGCTT----------ATTTTTAATTTTTTGTTGTTTGAAATGGAAT 930
         |||||||| |||||           |||||||||||||||||||||||||| |||
Db  2318 CACCACATCTGGCCTGTTGATATTTTAATTTTTTTTTTTGAGATAGAGT 2377

Qy   931 CTCGCTCTGTCGCCCAGGCTGGGAGTGCAATGGCCAACCCCGT 990
         |||||| |||| ||||||||||||||||||||||||||||||||
Db  2378 CTCACTCTGTTACCCAGGCTGGAGTGCAATGGCAGTGGTCGCAGCTCACTGCAACCTCCGT 2437
```

RESULT 13

```
Qy   991  CTCCCGGGCTCAAGCGATTCTCCTGTCTCAGCCTCCCAAGCAGCTGGGATTACGGGCACC  1050
          |||||| || |||| ||||| ||||||||||||||||||  |||||||||||| |  |
Db  2438  CTCCCGGGTTCAGGTGATTTTCCTGTGTCAGCCTCCCGAGTAGCTGGGATTACAGGTGTC  2497

Qy  1051  TGCCACCACACCCCGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATATTTGT  1110
          | |||||||| | ||||||||||||||||||||||||||||||||||||||||||  | |
Db  2498  CACCACCAAGCCTGGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGC  2557

Qy  1111  CAGGCTGGTCTCAAAACTCCTGACCTCAGGTGATCCACCCGCCTCGGCCTCCCAAAAGTGCT  1170
          ||||||||||| ||| ||||||||||   |||||  | |||||||||||||||||| |||||
Db  2558  CAGGCTGGTCTCGAACTCCTGACCTC--GTGATCTGCCTGCCTCGGCCTCCCAAAGTGCT  2615

Qy  1171  GGGATTACAGGCGTGAGCCACCCACCTCACCCACCGGGCTCAAATTTAGAT  1215
          |||||||||||  |||||||||||||   ||||||||||| |  || | ||
Db  2616  GGGATTACAGGTGTGAGCCACCCACCGTGCCCACCGGGCTTATTTATAT  2660
```

LOCUS   AV700988   839 bp   mRNA   linear   EST 16-JAN-2002
DEFINITION AV700988 GKC Homo sapiens cDNA clone GKCAEA01 3', mRNA sequence.
ACCESSION   AV700988
VERSION   AV700988.1 GI:10302959
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 839)
  AUTHORS   Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
            Xiao,H., Qu,J., Liu,F., Huang,Q., Cheng,Z., Li,N., Du,J., Hu,W.,
            Shen,K., Lu,G., Fu,G., Zhong,M., Xu,S., Gu,W., Huang,W., Zhao,X.,
            Hu,G., Gu,J., Chen,Z., and Han,Z.
  TITLE   Insight into hepatocellular carcinogenesis at transcriptome level
            by comparing gene expression profiles of hepatocellular carcinoma
            with those of corresponding noncancerous liver
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
  MEDLINE   21625106
COMMENT   Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shouing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES   Location/Qualifiers
     source   1..839
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="GKCAEA01"
                     /clone_lib="GKC"
                     /tissue_type="hepatocellular carcinoma"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
BASE COUNT   144 a   259 c   203 g   231 t   2 others
ORIGIN

Query Match   25.4%; Score 366.6; DB 10; Length 839;
Best Local Similarity   81.4%; Pred. No. 5.5e-58;
Matches 499; Conservative 0; Mismatches 101; Indels 13; Gaps 6;

```
Qy   602  TTTTTTATTTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGTG  661
          | | |||| |||||||||||||||||  ||||||| ||||| |||||||||||||||||| 
Db    66  TTCTTTTTTTTATTTTTTGAGCAGGAGTCT-GGCTCGTGTCACCAGGCTGGAGTGCAGG  124

Qy   662  GCGCAAATCTTGGCCTCACTGCAACCTCCGCCTCCCGGGTTCAAGTTATTCTCTCGCCCCAG  721
          |||| | ||| || |||||||||| ||||  | ||||||||| |  | |||| ||||| ||
Db   125  GCGCGATCTCGGCTCACTGCAAGCTCCGTTTCCTGGGTTCACCCATTCTCCTGCCTCAA  184
```

RESULT 14

```
Qy   722  CCTCCTGAGTGAGCTGGGACTACAGGCG---CCCACCACGCCCGGCTAGCTAATTTTTGTATT  778
          ||||| ||| | |||| ||| ||||||||   |||||||||||| |||||||||||||| |||
Db   185  CCTCCCGAGTAGCTGTGAATACAGGCGTTCGCCACCACGCCCGACTAATTTTTTGTATT  244

Qy   779  TTTAGTAGAGATGGGG--TTCACCATGTTCGCCAGGCTGGTCTCGATCTCCTGACCTTGTG  837
          ||||||||||| ||||  |||||| |||| ||||||| |||| |||||||||||||| |||
Db   245  TTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTGTCGATCTCCTGACCTCGTG  304

Qy   838  ANCTGGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCACACGCGCCCGGC  897
          | ||| |||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db   305  ATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCGCCCGGC  364

Qy   898  TTATTTTTTAATTTTTGTTTGTTTGAAATGGAATCTCACTCTGTTACCCAGGCTGGAGTGC  957
          | | ||||| |||||| ||||| |||| ||||| |||||| |||| |  | |||||||||||
Db   365  CTCT------TATTTTTTTTTTTTTGAGATGGAGTCTCACACTGTCACACGCTGGAGTGC  418

Qy   958  AATGGCACAAATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAAGCGATTCTCCTGCC  1017
          | ||| |   |||||||||||||||||||||| |||||||||||||||  ||||||||||
Db   419  AGTGGAGGGATCTCGGCTCACTGCAACCTCCGNCTCCCGGGTTCAAGAGATTCTCCTGCC  478

Qy  1018  TCAGCC-TCCCAAGCCAGCTGGGATTACGGGCACCGCCACCACCACCCCGGCTA-ATTTTTG  1075
          ||||||  |||||| |||||||||||| |||  ||  |||||| | |||||||  ||||||
Db   479  TCAGCCTTCCCAAGTAGCTGGGATTACAGGTGCCCACCACCACGCCCGGCTAGTTTTTTG  538

Qy  1076  TATTTTCATTAGAGAGGCGGGGGTTTCACCATATTTGTCAGGCTGGTCTCAAAACTCCTGACCT  1135
          ||||||  | |||| ||| |||| ||||||| | | || |||||||||| |||||||||||| |
Db   539  TATTTTTAGGCAAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACAT  598

Qy  1136  CAGGTGACCCACCTGCCTCAGCCTTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCACCTCA  1195
          |||||||  | || | || ||  | ||||| |||||||||||||||||||||||||||||| |
Db   599  CAGGTGAACCGCCCACCTTAGNCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACCACCATA  658

Qy  1196  CCCAGCCGCGGCTAA  1208
          ||  ||||||  |||
Db   659  CCTGGCCGCAAAA  671
```

LOCUS   BM722145   671 bp   mRNA   linear   EST 01-MAR-2002
DEFINITION UI-E-E00-ahy-e-13-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone
            UI-E-E00-ahy-e-13-0-UI 5', mRNA sequence.
ACCESSION   BM722145
VERSION   BM722145.1 GI:19042520
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 671)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE   Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT   Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 5242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr Gregg Hageman
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            The following repetitive elements were found in this cDNA
            sequence: 39-327, >ALU (matched compliment) 352-640, >ALU (matched
            compliment)
            Seq primer: M13 Reverse.
FEATURES   Location/Qualifiers

First alignment block

```
source          1..671
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="UI-E-EO0-ahy-e-13-0-UI"
                /clone_lib="UI-E-EO0"
                /tissue_type="fetal eye"
                /dev_stage="fetal"
                /lab_host="DH10B (Life Technologies)"
                /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                modified polylinker; Site_1: EcoR I; Site_2: Not I;
                UI-E-EO0 is a cDNA library containing the following
                tissue(s): fetal eye. The library was constructed
                according to Bonaldo, Lennon and Soares, Genome Research,
                6:791-806, 1996. First strand cDNA synthesis was primed
                with an oligo-dT primer containing a Not I site. Double
                stranded cDNA was ligated to an EcoR I adaptor, digested
                with Not I, and cloned directionally into pT7T3-Pac
                vector. The oligonucleotide used to prime the synthesis of
                first-strand cDNA contains a library tag sequence that is
                located between the Not I site and the (dT)18 tail. The
                sequence tag for this library is CGCGTATACC. This library
                was created for the program, Gene Discovery in the Visual
                System, supported by National Eye Institute (NEI)."

BASE COUNT     113 a     208 c     165 g     185 t
ORIGIN

Query Match
Best Local Similarity   79.4%;
Matches  510;  Conservative   0;  Mismatches  115;  Indels  17;  Gaps   6;
                                Score 364;  DB 14;   Length 671;
                                Pred. No. 1.8e-57;
```

```
Qy    597 TTTATTTTATTTATTTTTTAATTTTTTTTAATTTTTTTTGAGACAGAGTCTCAACTCTGTCA 656
          |||||| ||||||||||||| |||||||||||||||||||||| |||||||||||||||||
Db     24 TATACCTTTTTTTTTTTTTTTATTTTTTTTTTTTTTGAGACTGAGTCTC-GCTCTGTTCGCCCAGGCTGGAGTG 82

Qy    657 CAGTGGCGCAATCTTGGCTCACTGCAACCTCTCGCCTCCCGGGTTCAAGTTATTCTCCTGC 716
          |||||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||
Db     83 CAGTGGCGCGATCTCGGCTCACTGCAAGCTCCGCCTCGGGGTTCACGCCATTCTCCTGC 142

Qy    717 CCCAGCCTCCTGAGTAGCTGGGACTACAGGCG---CCCACCACGCCCTAGCTAATTTTTTT 773
          ||||||||| ||||||||||||||||||||        ||||||||| |||||||||||
Db    143 CTCAGCCTCCCAAAGTAGCTCGGACTACAGGCGTCCTCCACCATCGCTAATATTTTTG 202

Qy    774 GTATTTTTAGTAGAGAGATGGGGG-TTCACCATGGTTCGGCCAGGTTGGATCTCTGGACC 832
          ||||||||||| ||||||||||||| ||||||||||| ||||| |||||||||||||||
Db    203 TAATTTTTTAGTAGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTCTCGATCTCCTAACC 262

Qy    833 TTGTGTGATCTCGGCCTCTGCCTCGGCCCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGC 892
          ||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    263 TGGTGATCCGCGCCGCCTCCAAAGTATTGGGATTACACCGCGAGCCACCGCG 322

Qy    893 CCGGC------------TTATTTTAATTTTTTGTTTGTTTTGTTTTGAAATGGAAATCTCACTCTGTTAC 943
          |||||           |||||||||||||||||||||||||||||||||| |||||||||
Db    323 CCGGCTGCAAAATAAATCTTCTTTTTTTTTCTTTTTTCTTTTTTTGTGAGACAGAGTCTCGCTCTGTTGC 382

Qy    944 CCAGGCTGGAGTGCAATGGCAAAATCGCGCTCACTCCAACCTCTCGCCTCCCGGGTTCAA 1003
          |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    383 CCAGGCTGGAGTGCAGTGGCGACGATCTCGGCTCACTGCAACCTCCGCCTCCCGGGTTCAC 442

Qy   1004 GCGATTCTCCTGTGTCAGCCTCCCGAAGCAGCTGGGATTACGGGCGCACCGACCACCACC 1063
          |||||||||| |||| ||||||||||| ||||||||| |||||||||||| |||||||||
Db    443 GCCATTCTCGTCGCCCCAGCGTTCCCAGGAGCTAGCTGGGACTACAGGCGCCCCGCCACCGCC 502

Qy   1064 CGCTAA-TTTTTGTATTTTCATTAGAGGGGGGGTTTCACCATTATTTGTCAGGCTGGTCTC 1122
          |||||| ||| || |||||||||||||||||||||||||||||| |||||||||||| |||
Db    503 GGCTAACTTTTTGTGTGTTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGATGGTCTC 562

Qy   1123 AAACTCCTGACCTCAGGTGACCCACCTGCCTCCCAAAGTGCTGGGATTACAGGC 1182
          ||||||||||||||||||| |||||||| | ||||||||||||||||||||||||
Db    563 GATCTCCTGACCT--TGTGATCTGCCCGCCTCAGCCTGCCAAAGTGCTGGGATTACAGGC 620

Qy   1183 GTGAGCCACCACGCCCGGGCGGCTAATTTTAGATATAAAAAAAATATA 1224
          |||||||||||||| |||||| | | |||||||
Db    621 GTGAGCCACCACACCAGGCCCGAGAAATGCTTTTTTTTAAAAA 662
```

Second alignment block

```
RESULT 15
LOCUS          BM996878                    611 bp     mRNA     linear     EST 17-JUN-2002
DEFINITION     UI-H-DH0-aun-n-14-0-UI.s1 NCI_CGAP_DH0 Homo sapiens cDNA clone
               IMAGE:5871949 3', mRNA sequence.
ACCESSION      BM996878
VERSION        BM996878.1   GI:19721779
KEYWORDS       EST.
SOURCE         human.
  ORGANISM     Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1  (bases 1 to 611)
  AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
  JOURNAL      Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Dr. Jose Mercuende
               cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
               cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
               DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
               Clone Distribution: Clone distribution information can be found
               through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
               The following repetitive elements were found in this cDNA
               sequence: 11-297, >ALU (matched compliment) 307-595, >ALU (matched
               compliment)
               Seq primer: M13 FORWARD
               POLYA=Yes.
FEATURES       Location/Qualifiers
     source    1..611
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:5871949"
               /clone_lib="NCI_CGAP_DH0"
               /tissue_type="Metastatic Chondrosarcoma"
               /dev_stage="Adult"
               /lab_host="DH10B (Life Technologies)"
               /note="Organ: lung; Vector: pT7T3-Pac (Pharmacia) with a
               modified polylinker; Site_1: EcoR I; Site_2: Not I;
               NCI_CGAP_DH0 is a cDNA library containing the following
               tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma
               in Lung. The library was constructed according to Bonaldo,
               Lennon and Soares, Genome Research, 6:791-806, 1996. First
               strand cDNA synthesis was primed with an oligo-dT primer
               containing a Not I site. Double stranded cDNA was ligated
               to an EcoR I adaptor, digested with Not I, and cloned
               directionally into pT7T3-Pac vector. The oligonucleotide
               used to prime the synthesis of first-strand cDNA contains
               a library tag sequence that is located between the Not I
               site and the (dT)18 tail. The sequence tag for this
               library is AGATCATTGC.
               TAG_LIB=UI-H-DH0
               TAG_TISSUE=lung
               TAG_SEQ=AGATCATTGC"

BASE COUNT     119 a     173 c     139 g     180 t
ORIGIN

Query Match
Best Local Similarity   79.4%;
Matches  479;  Conservative   0;  Mismatches  116;  Indels   8;  Gaps   4;
                                Score 361.4;  DB 14;   Length 611;
                                Pred. No. 5.6e-57;
```

```
Qy    604 TTATTTTTAATTTTTTGAGACAGAGTCTCAACTCTGTCACCCAGGCTGGAGTGCAGTGGC 663
          |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTTTTTTTTTTTTTTTTTTTTGAGATGGAGTCTTGCTCTATAGCCCAGGCTGGAGTGCAGTGGC 60

Qy    664 GCAATCTTGGGCTCACTGCAACCTCTGCCTCCCGGGTTCAAGTTATTTCTCCTGCCCCAGCC 723
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||| ||||
Db     61 ACAATCTTGGCTCACTGCAACCTCCGCCTCCCAGTTCAAGTGAATTCTCCTGCCTCAGCC 120
```

```
Qy    724 TCCTGAGTAGCTGGGACTACAGGGCGCCCACCACGGCCTAGC-TAATTTTTTGTATTTTTA 782
          ||||||||||||||||||||||||||||| ||||| ||| |  | |  |||||| ||||||
Db    121 TCCCGAGTAGCTGGGACTACAGGGTGCGTGCTACCACGCCCATCTAATTTTTGTATTTTTA 180

Qy    783 GTAGAGATGGGG-TTCACCATGTTTCGCCAGGTTGATCTTGGACCTTGTGATCT 841
          |||  ||||||| ||||||||||||||||||| ||||||||||| |  ||||||
Db    181 GTACAGATGGGGTTTCACCCTGTTTGGCCAGGATGGTCTCGATTCTCTTGACTTCGTGATCC 240

Qy    842 GCCTGCCTCGGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGCTTAT 901
          || || |||||||||||| | |||||| |||||||| ||| |||  ||||||||| || |
Db    241 ACCCACTTCGGGCCTCCCAAAAGTGCTGGGATTACAGGCGTGAGCCACCACGCCCGGC----- 296

Qy    902 TTTTTAATTTTTTGTTTTGTTTTTGAAATGGAAATCTCACTCTGTTACCCAGGCTGGAGTGCAAATG 961
          ||| | | |  | || |  | ||||||||||||||||||||||||||||||||||||||||||||
Db    297 -CGCATTTCTTTCTTTTTTTTGAGACAGGGTCTCACTCTGTCACCCAGGCTAGAGTGCAAAG 355

Qy    962 GCCAAAATCTCGGGCTCACTGCAAACCTCTGCCTTCCCGGGCTCAAGCGATTCTCCTGTCTCCAG 1021
          ||||||| |||||| ||||||||||| |  ||  | |||||||||||||||||||| | |||||
Db    356 GCGC-ATCTCGGCTCACTGCAACGTCCGCCTTCCTGGGTTCAAGCGATTCTCATCCTTCAG 414

Qy   1022 CCTCCCAAGCAGCTGGGATTACGGGCACCTGCCACCACCCCGCTAATTTTTTGTATTTT 1081
          || || ||||||||| ||||||||||| | |||| | | |||||||||||||||||||
Db    415 CCACCTGAGTAATTGGGATTACAGGCGCACATGCCACCACGCCCAGCTAATTTTTTGTATTTTT 474

Qy   1082 CATTTAGAGGCCGGGGGTTTCACCATATATTTGTCAGGCTGGTCTCAAAACTCCTGACCTCAGGTG 1141
          | | ||||| || ||||||||||||||| ||  |||||| || |||||||| ||||| |||||||
Db    475 TAGTAGAGACGAAGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCATGACCTCAAGTG 534

Qy   1142 ACCCACCTGCCTCAGCCTTCCAAAAGTGCTGGGATTACAGGGGTGAGCCACCTCACCCAGC 1201
          ||||| ||| |||||| ||||||||||||||||||||||| || |||||||| ||||||
Db    535 ATCCGCCCGCCTTCGGCCTCCCAAAAGTTTTGGGATTACAGGCATGAGCCACTACGTCCAGC 594

Qy   1202 CGG 1204
          | |
Db    595 CTG 597
```

Search completed: February 2, 2003, 00:27:28
Job time : 2234 secs